Structure-function relationship of the Epidermal Growth Factor Receptor on the cell surface

A microscopy study using fluorescent inhibitors

Korovesis, Dimitrios

Awarding institution:
King's College London

Structure-function relationship of the Epidermal Growth Factor Receptor on the cell surface: A microscopy study using fluorescent inhibitors

A thesis submitted to the Kings College London for the degree of Doctor of Philosophy

by

Dimitrios Korovesis

Division of Cancer Studies

Faculty of Life Sciences & Medicine

King's College London

2017

This thesis is dedicated to my other half, Lena, who has been the driving force for personal development and for her inexhaustible support, faith and patience even in the "darkest" days during my PhD

Declaration

I hereby declare that the work presented in this thesis is the result of my own investigations and has not been accepted in a previous application for a degree.

Acknowledgements

I would like to express my gratitude to my supervisors Dr Simon Ameer-Beg and Dr Marisa Martin-Fernandez for giving me the chance to be a part of King's College London and Central Laser Facility and also for their helpful advice, guidance, patience and support throughout my studies.

I would also like to thank all members of Dr Marisa Martin-Fernandez group (FBI team of the LSF, Harwell) who helped to develop my research project. In particular, I am grateful to Dr Christopher J. Tynan, Dr Selene K. Roberts, Dr Laura Zanetti-Domingues, Dr Sarah R. Needham, Dr Daniel J. Rolfe and Dr Michael Hirsch for the great patience they showed, despite my initial lack of knowledge in the field of Biology and Imaging, and constant scientific and emotional support.

I am also thankful to Dr Martyn Winn and Dr Valeria Losasso for the molecular modelling and to Dr Stephen E. D. Webb and Mr Benjamin C. Coles for looking after me when I encountered technical problems with the microscopes.

I am thankful to Evotec AG for providing the facilities, reagents and services for the successful completion of the organic synthesis. I also wish to thank my collaborators in Evotec AG as without them the chemical synthesis would not have been possible. Especially I would like to thank Dr Joanna Lisztwan, Mr Tim Jones and Dr Robert Townsend for their constructive thoughts, suggestions, support and hospitality during my stay in Evotec AG.

Many thanks to: Dr Paul van Bergen en Henegouwen of Utrecht University for the gift EgB4 and 9G8 nanobodies, Prof Linda Pike of Washington University for the gift of a Cho cell line stably and inducibly expressing EGFR alleles, Dr John Löfblom of KTH Royal Institute of Technology for providing the ErbB3 Affibody plasmid and Dr D. Tomlinson of University of Leeds for providing anti-ErbB4 Adhiron.

Special thanks to my family Kostas, Amalia, Nikos and Giorgos, for their support throughout my life but also to my friends and fellow scientists Dr Giorgos Paraskevopoulos, Mr Harris Georgiou and Dr Kostas Papadopoulos for the encouragement and the useful discussions.

Last but definitely not least, I want to thank from the bottom of my heart the light of my life, my wife Lena Petraki, for being by my side and putting up with me during this work. Finally, my beloved dogs, Bijou and Ira, for their endless affection and unconditional loyalty despite the distance between us, which was the price I had to pay in order to pursue my dream to become a scientist.

Abstract

The Human Epidermal Growth Factor Receptor (EGFR or HER) family consists of four members (ErbB or HER1-4) and belongs to the super-family of receptor tyrosine kinase (RTK). Mutation and/or over-expression of these receptors are implicated in various cancer types and they have been subject of intense research for anti-cancer therapies. However, effectiveness of current targeted therapeutics, either treatment with tyrosine kinase inhibitors (TKIs) or antibodies (and even a combination of both) differs from patient to patient. Even in cases when these therapies are initially effective, the disease returns back exploiting its complexity (EGFR signalling through different pathways) or other defence mechanism (e.g. altered signalling via EGFR mutations). A better understanding of EGFR structure and TKIs mechanism of action are therefore imperative to understand the function of the receptor during treatment and for improving the efficiency of future therapeutics.

In order to address some of the aforementioned unknown aspects, I require tools and methods that would allow me to investigate the structure of the receptor on cells in basal and EGF activated states as well as the conformational changes induced by TKI binding.

Following the development of methods previously reported by my group to probe the structure of the ECD, I have developed a range of different fluorescently labelled TKIs (fTKIs) for labelling the ICD of the receptor *in vitro*. These fTKIs are comprised of the pharmacophore and of the fluorophore and target the ATP-binding site of receptor which is the "heart" of the ICD. Characterization of fTKIs' binding by confocal microscopy showed their ability to be used as reporters of receptor's activation state. For the ECD labelling, I have used existing fluorescent probes (non-activating molecules and activating ligands) and have employed a combination of advanced single-molecule methods (Single-Particle Tracking and Fluorescence Localization Imaging with Photobleaching) and FRET measurements (Distance Of Closest Approach). The current investigation revealed new information regarding the architecture and the oligomerization of EGFR in cells as well as insight into TKIs conformational preference and drug resistance mechanism.

Table of Contents

List of Figures

List of Schemes

List of Tables

Abbreviations

AcOH	acetonitrile
Affy	Affibody
ATP	adenosine triphosphate
Ar	aromatic
A.U.	arbitrary units
BSA	bovine serum albumin
^{t}Bu	*tert*-butyl
^{o}C	degrees Celcius
$CDCl_3$	deuterated chloroform
COSY	correlation spectroscopy
δ	chemical shift
d	doublet
Da	Dalton
dd	double doublet
ddd	double double doublet
DCM	dichloromethane
DEPT	distortionless enhancement by polarization Transfer
DI	domain I
DIII	domain III
DIPEA	*N*,*N*-diisopropylethylamine
DMAP	4-*N*,*N*-dimethylaminopyridine
DMEM	Dulbecco's modified Eagle's medium
DMF	*N*,*N*-dimethylformamide
DMSO	dimethyl sulfoxide

DOCA	distance of closest approach
ECD	extracellular domain
EM	electron microscopy
ESI	electrospray ionization
etc.	*et cetera*; and so forth
Et_2O	diethylether
EtOAc	ethyl acetate
EtOH	ethanol
FDA	food and drug administration
FLImP	fluorophore localization imaging with photobleaching
fTKI	fluorescent tyrosine kinase inhibitor
GA	glutaraldehyde
h	hour(s)
HCl	hydrochloride
HMQC	heteronuclear multiple quantum coherence
HPLC	high perfomance liquid chromatography
HRMS	high resolution mass spectrometry
Hz	Hertz
ICD	intracellular domain
in vivo	in living organism
in vitro	in glass
in situ	in place
i-PrOH	isopropanol
J	spin-spin coupling constant
JM	juxtamembrane domain

K_d	dissociation constant
K_{off}	dissociation rate
K_{on}	association rate
LC-MS	liquid chromatography mass spectrometry
m	multiplet
MD	molecular dynamics
Me	methyl
MeCN	acetonitrile
MeOH	methanol
$MgSO_4$	magnesium sulphate
MHz	megahertz
min	minute(s)
m/z	mass to charge ratio
MS	mass spectrometry
MTBE	methyl tert-butyl ether
NMR	nuclear magnetic resonance
NMP	n-methyl-2-pyrrolidone
PDB	protein databank
PEG	polyethylene glycol
PFA	paraformaldehyde
$POCl_3$	phosphoryl chloride
ppm	parts per million
pTsOH	p-toluenesulfonic acid
q	quartet
R_f	retention factor
R_t	retention time

s	singlet
SNR	signal-to-noise
SPT	single-particle tracking
t	triplet
t-BOC	*N-tert*-butoxycarbonyl
tert	tertiary
THF	tetrahydrofuran
TKD	tyrosine kinase domain
TKI	tyrosine kinase inhibitor
TLC	thin layer chromatography
TM	transmembrane domain
TMR	tetramethylrhodamine
WB	western blot
wt	wild type

1 General Introduction

1.1 Receptor Tyrosine Kinases

Many important cell processes such as proliferation, survival, differentiation, function and motility rely on the communication of the extracellular with the intracellular environment. The transduction of the information across the plasma membrane is performed by cell surface receptors. Amongst other receptors, the Receptor Tyrosine Kinase (RTK) super-family has been associated with many of the aforementioned processes in both health and disease[1].

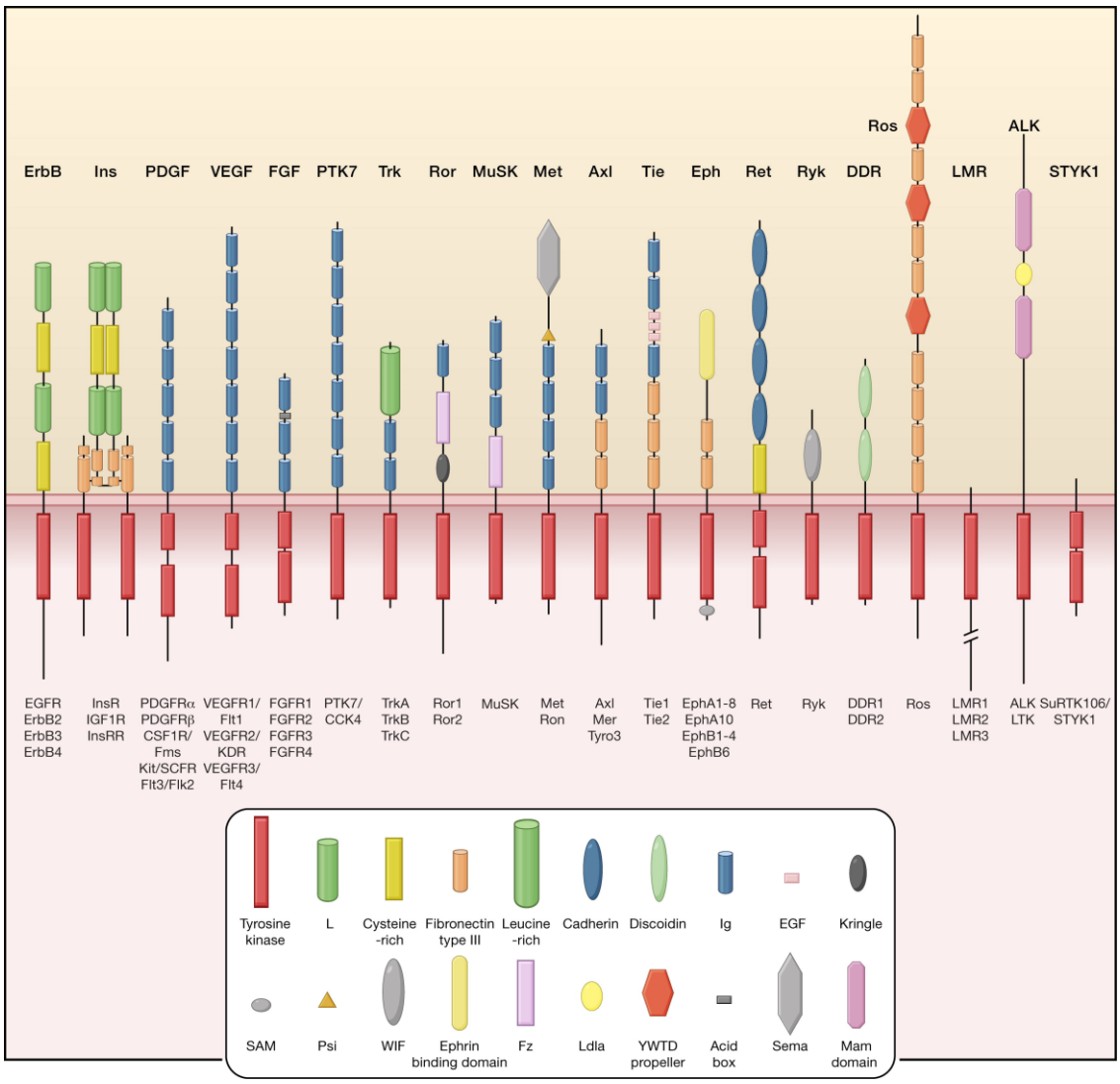


Figure 1-1: Receptor Tyrosine Kinase Families. All RTKs consist of an extracellular and an intracellular kinase domain. In all cases, receptor activation requires growth factor binding which leads to C-tail phosphorylation of the dimers-oligomers (Adapted from M. A. Lemmon and J. Schlessinger, *Cell*, 2010)[1].

The RTK superfamily consists of 58 proteins which can be classified into 20 sub-families as illustrated in ***Figure 1-1***. All of receptors share the same structural architecture consisting of an extracellular domain (ECD), a transmembrane segment (TM) and an intracellular domain (ICD) with kinase activity. The signal transducing function of most RTKs in healthy cells depends on stimulation by growth factors which bind to an epitope on the ECD. Although dimerization of RTKs was first thought to be a common step during activation, more recent studies have shown that some receptors exist as pre-dimers[2]. Moreover, higher order oligomers have also been proposed as important in signalling of some RTKs[3,4]. Despite the differences in organization on the cell surface in the basal state, an activated signalling unit requires non-monomeric ligand-bound receptors. Phosphorylation of RTKs allows activation of different downstream signalling pathways for regulating cell processes. The first characterised protein of the RTK super-family was the Epidermal Growth Factor Receptor (EGFR).

1.2 The EGFR Family of Receptors: History, Importance and Current Therapeutics

The history of the EGFR family of receptors started more than 50 years ago in 1962 when Stanley Cohen isolated and characterised the protein responsible for inducing anatomical changes in new-born mice and rats[5]. This finding was a result of a research which had started a few years earlier and led to the discovery of the nerve growth factor (NRG) by Cohen and Levi-Montalcini[6] which also initiated the research on growth factors. The isolated protein was named epidermal growth factor (EGF) due to its ability to stimulate epidermal proliferation[7]. The EGF receptor (or ErbB1 or HER1) was first reported in 1980 by Cohen *et al.*[8] after investigations to find the target of the EGF ligand by using a radiolabelled EGF molecule (^{125}I-EGF). Subsequently after Cohen's discovery, the identification of the rest of the family members (ErbB or HER 2-4) was followed by different groups[9]. Since then, these receptors have been the target of a significant amount of cancer research due to their identification in different cancers[10] and have become one of the most studied RTK systems.

The EGFR family consists of 4 homologous receptors with the same structural (ECD, TM and ICD) and functional features (activation through ligand binding and dimerization-oligomerization) mentioned previously for RTKs (***Figure 1-2***). To date, approximately 13 ligands are known to bind and regulate ErbB family members signalling[11], however none of them bind ErbB2 (also referred as *neu*). ErbB3 is also distinct from the others as it is catalytically inactive or has low activity, even though it is competent to bind ATP[12]. As a consequence, activation of these two receptors occurs only through hetero-dimerization with the other members of the family as well as with each other.

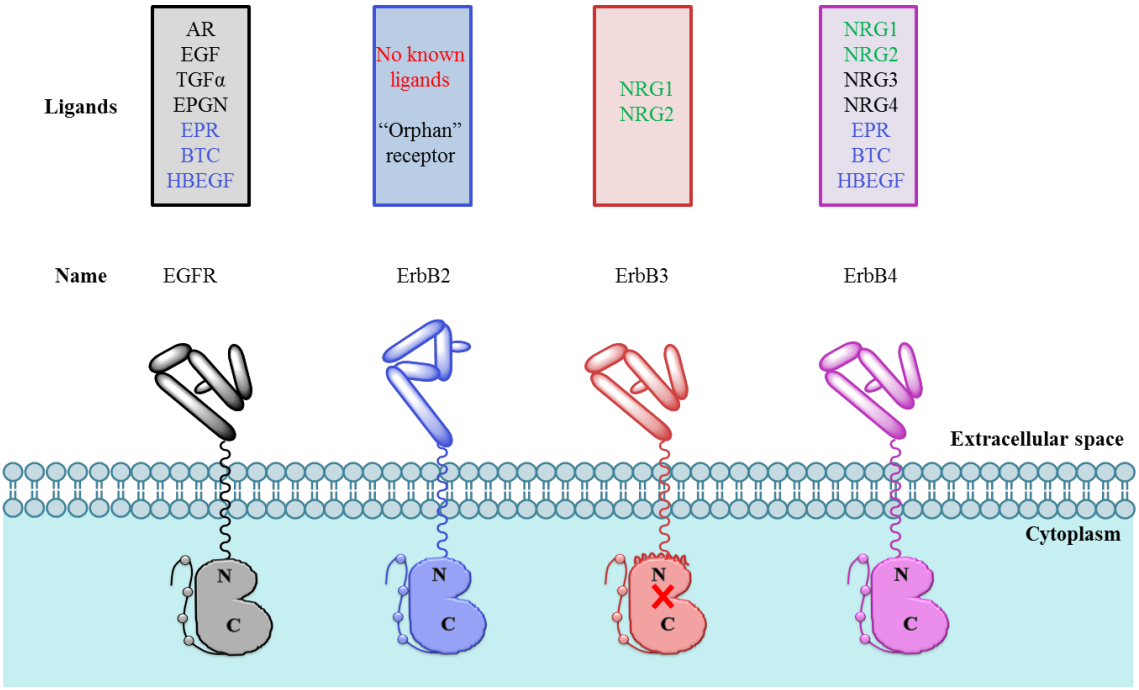


Figure 1-2: The ErbB family members and their ligands. All EGFR members share a similar architecture consisting of an extracellular domain, a helical transmembrane domain, and an intracellular domain. The main differences between them are that there are no known ligands for ErbB2 and that ErbB3 is catalytically inactive (red X on the kinase domain). Common ligands between EGFR and ErbB4 receptors have blue coloured letters and between ErbB3 and ErbB4 with green. Ligands coloured with black (EGFR and ErbB4) are specific to the respective receptors.

Activation through homo- or hetero-dimerization of ErbBs leads to tyrosine phosphorylation of their C-terminal tails which are being used as docking sites for downstream effectors with Src Homology 2 (SH2) or PhosphoTyrosine Binding (PTB) domain[13]. Association of these effector molecules activates a complex signalling network which regulates a variety of essential biological processes. The most common

signalling pathways are: the Mitogen-Activated Protein (MAP) kinase cascade, the Phosphatidylinositol 3-kinase/protein kinase B (PI3K/AKT), the Phospholipase-Cγ (PLCγ) and the Signal Transducers and Activators of Transcription (STAT) pathways (***Figure 1-3***).

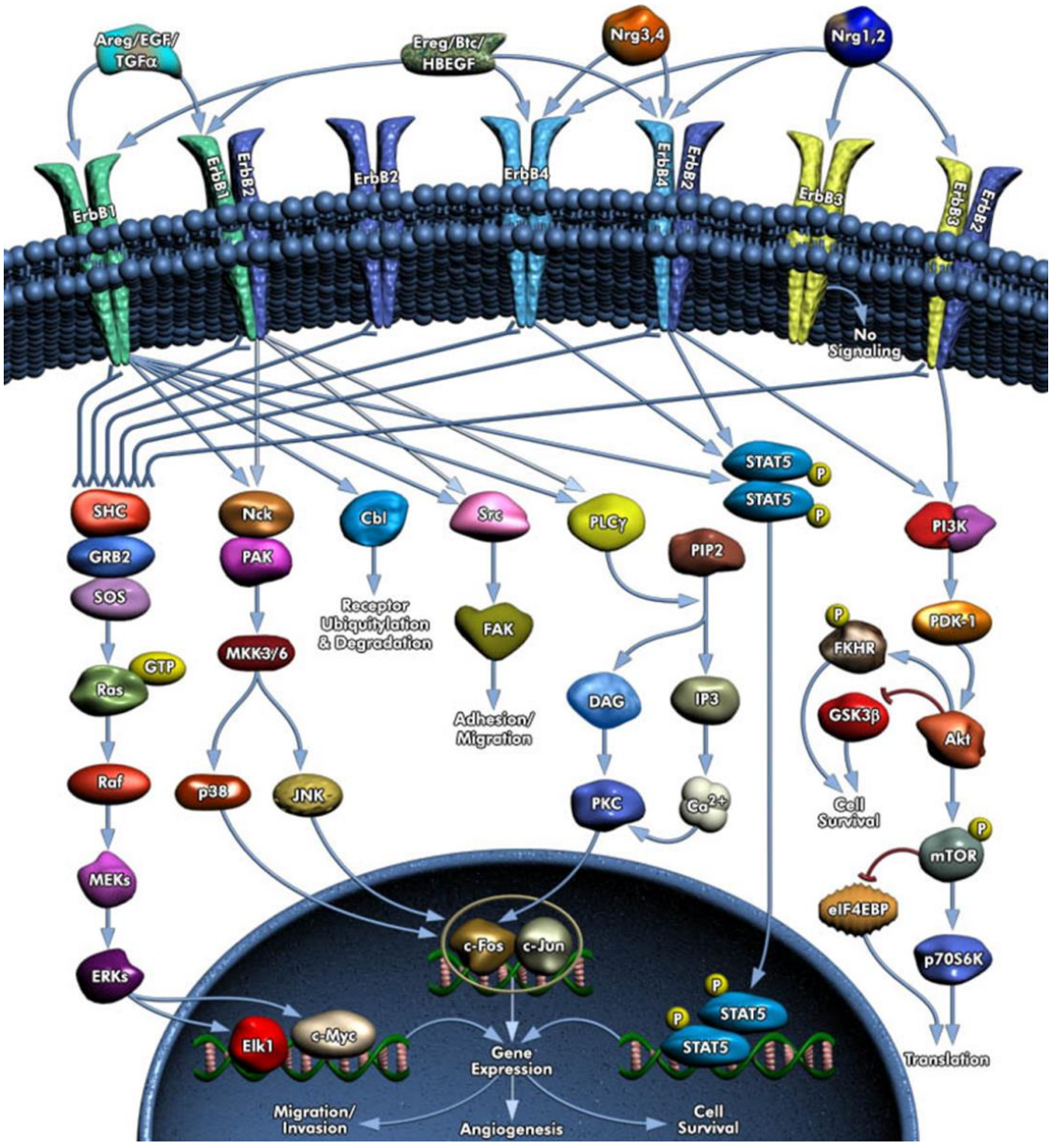


Figure 1-3: The most common signalling pathways of the EGFR family. Phosphorylation of ErbB homo- and hetero-dimers results in the recruitment of downstream effectors which propagate the signal through different pathways such as MAP, PI3K/AKT, PLCγ and STAT (Adapted from Qiagen website).

The ErbB family of receptors plays a prominent role in many cancer histotypes, expect from their physiological roles in cells (regulate apoptosis, cell cycle progression, differentiation, development, and transcription). The association of ErbB family of

receptors in the pathogenesis of cancer was first reported in 1985 when Thompson et al.[14] provided evidence for the correlation of EGFR overexpression and cancer. This correlation also consists the first example for a relationship between receptor overexpression and cancer[15]. Since then, dysregulation of the physiological function of these receptors either by mutation or gene amplification (or both simultaneously) is implicated in the development of many types of cancer including lung, breast, stomach, colorectal, ovary, head and neck, pancreas, and pancreatic carcinomas and glioblastoma[9,15].

For these reasons, ErbBs are key targets of cancer therapies. The most used approaches in clinic are treatment with either Tyrosine Kinase Inhibitors (TKIs)[16] or monoclonal Antibodies (mAbs)[17] and in some cases where a patient develops resistance, a combination of both[18–21]. TKIs are small-molecule drugs that target the Tyrosine Kinase Domain (TKD) of a protein and block its phosphorylation by competing with ATP, thereby preventing the downstream signalling cascade[22]. On the other hand, mAbs bind to the ECD and their mode of action for, indirectly, blocking receptor phosphorylation varies. Some mAbs compete with and block ligand (e.g. EGF) binding (and subsequent dimerization)[23] whereas others prevent dimerization by "locking" the receptor in the inactive conformation[24]. ***Table 1-1*** below summarizes the approved drugs as well as some targeted therapeutics currently in clinical trials.

Table 1-1: Table of ErbB family members targeted therapeutics.

Therapy	Trade name	Target	Cancer Type	Status
		Monoclonal Antibodies		
Cetuximab	Erbitux®	EGFR	Colon and Head & Neck cancers	Approved
Panitumumab	Vectibix®	EGFR	Colon and Head & Neck cancers	Approved
Nimotuzumab	-	EGFR	Glioma, pancreatic, SCC	Approved
Necitumumab	Portrazza®	EGFR	NSCLC	Approved
mAb806	-	Δ2–7EGFR, EGFRvIII	Advanced solid tumors	Clinical trials
Trastuzumab	Herceptin®	ErbB2	Breast cancer	Approved
Pertuzumab	Perjeta®	ErbB2	Breast cancer	Approved

Seribantumab	-	ErbB3	Heregulin positive NSCLC	Clinical trials
Tyrosine Kinase Inhibitors				
Gefitinib	Iressa®	EGFR	NSCLC	Approved
Erlotinib	Tarceva®	EGFR	NSCLC	Approved
Lapatinib	Tykerb®	EGFR, ErbB2	Breast cancer	Approved
Afatinib	Giotrif®	EGFR, ErbB2	NSCLC	Approved
Dacomitinib	-	EGFR, ErbB2	SCC, NSCLC	Clinical trials
Sapitinib	-	EGFR, ErbB2, ErbB3	Various types	Clinical trials
Neratinib	-	EGFR, ErbB2, ErbB3, ErbB4	Breast cancer	Clinical trials
Osimertinib	Tagrisso®	T790M EGFR	T790M positive NSCLC	Approved
BI 1482694	-	T790M EGFR	T790M positive NSCLC	Approved
Rociletinib	-	T790M EGFR	T790M positive NSCLC	Clinical trials

NSCLC: *Non-Small Cell Lung Cancer,* ***SCC:*** *Squamous Cell Carcinoma*

Despite the development of various anticancer therapeutics for all ErbB members, their effectiveness differs from patient to patient. Even in cases when these therapies are initially effective, the disease returns back exploiting mainly resistance mechanisms such as somatic mutations on ErbBs' TKD[25–28] and altered downstream signalling thourgh homo- and heterodimerizations[29–32]. Improving the efficiency of future therapeutics is therefore tighly associated with our understanding of both the mechanism of activation and drug resistance of these receptors at a molecular level. Such an understanding could potentially allow scientists to better design drugs (either TKIs or mAbs) which would more accurately and efficiently inhibit ErbB members' activation.

1.2.1 Structural Analysis of ErbBs

ErbB receptors are type I transmembrane glycoproteins and each member consists of an ECD, a helical TM, a juxtamembrane segment (JM), a TKD and the C-terminal tail as illustrated in ***Figure 1-4***.

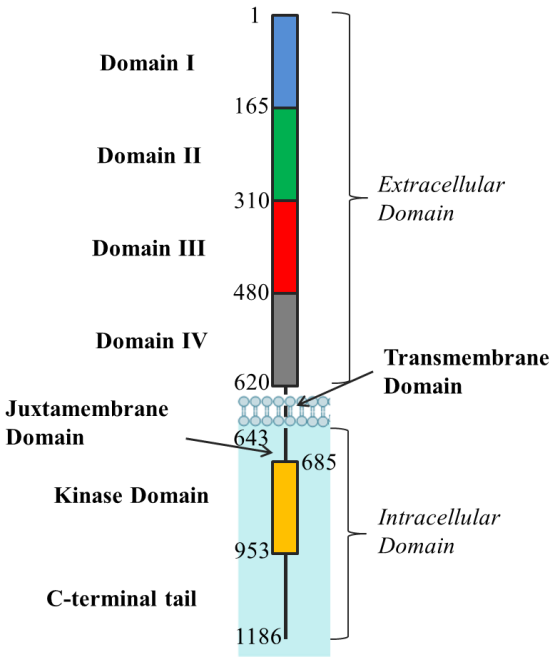


Figure 1-4: Structure of ErbB receptors. The ECD can be divided into four sub-domains (I-IV). Domains I and III are involved in ligand binding and domain II, which contains the dimerization arm, in dimerization. The TM, JM and TKD are also involved in the dimerization whereas the phosphorylated C-tail is served as a docking site for downstream signalling (*the amino acid numbering of each region is referred to EGFR*).

The **ECD** consists of approximately 600 amino acids and can be further divided into four sub-domains (I-IV): the leucine-rich sub-domains I (L1) and III (L2) and the cysteine-rich sub-domains II (CR1) and IV (CR2). Even though an early study, using titration calorimetry and X-ray scattering, had suggested that the formation of the ECD dimer occurs through direct interaction of two EGFR molecules[33], deeper insights into the active dimeric ECD form as well as the inactive-active transition were obtained a few years later from crystal structures[34–36]. In the inactive state, the receptor exists in a tethered-like state (***Figure 1-5A***) where CR1 and CR2 interact with each other in an auto-inhibitory conformation[36]. Ligand binding to the epitope between sub-domains I and III triggers conformational changes that interrupt the auto-inhibited interactions and allow dimerization by exposing the dimerization arm (sub-domain II)[34,35] as is shown in ***Figure 1-5B***. Both crystal structures of ECD dimers revealed a 2:2 ligand:ECD complexes (one structure with TGFα[34] and the other with EGF[35]) in a back-to-back configuration (***Figure 1-5C***).

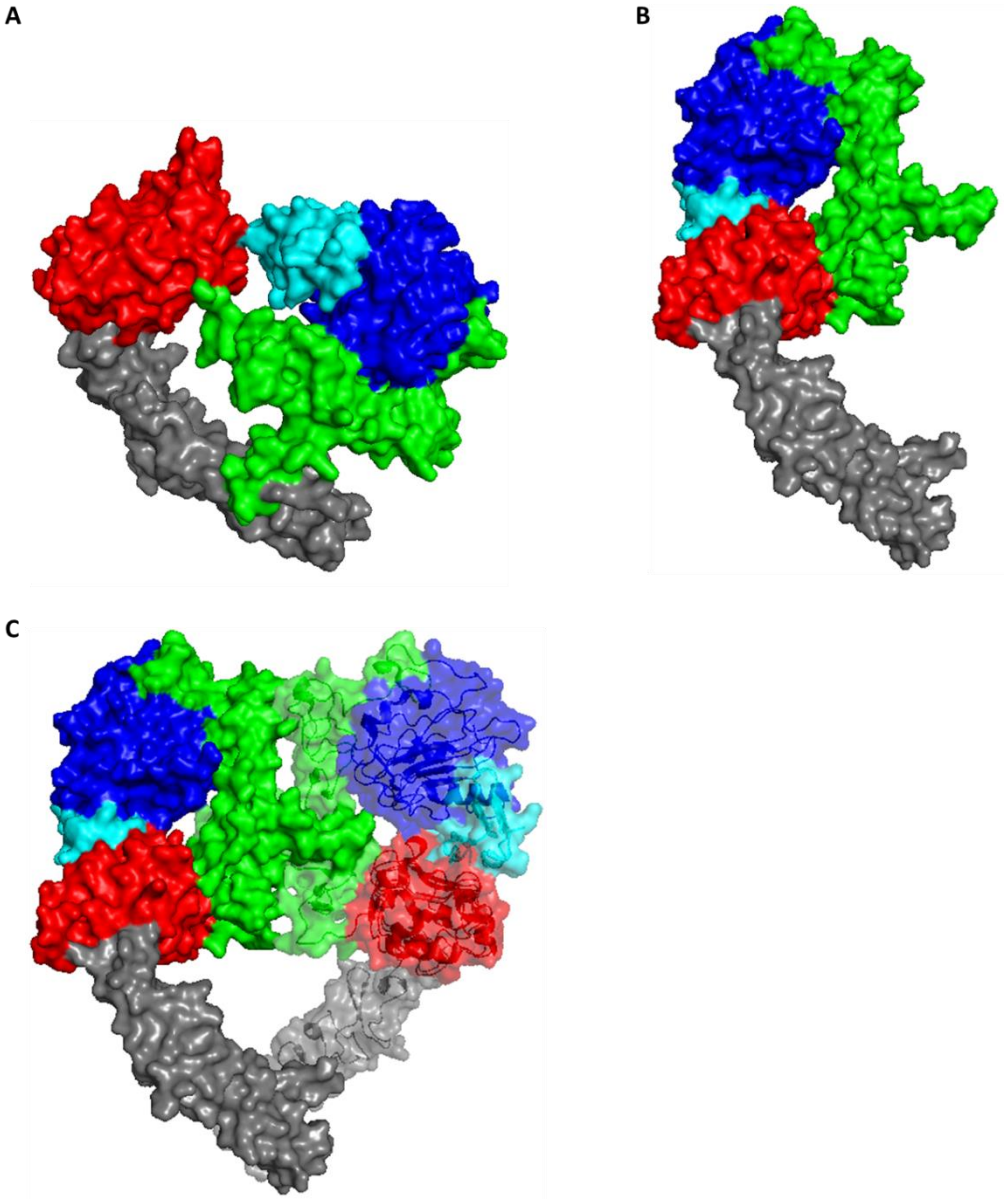


Figure 1-5: Crystal structures of the tethered and monomeric and dimeric extended conformation of the ECD of EGFR. EGF-EGFR complex of the tethered state (**A**, PDB: 1NQL) and the extended state (**B**, PDB: 3NJP). In **C** the extended ECD dimer of EGFR is depicted (PDB: 3NJP). *Colours: blue for sub-domain I, green for sub-domain II, red for sub-domain III, grey for sub-domain IV and cyan for EGF.*

Crystal structures of both ErbB3 and ErbB4 in the tethered conformation have also been resolved[37,38] and showed a similar configuration as in EGFR. The same it is not true for ErbB2 as the residues responsible for the intramolecular interaction between CR1 and CR2 are not conserved[39]. Therefore, the receptor always adopts the extended conformation[40,41]. This structure resembles the "active" extended conformation seen for EGFR and for ErbB4 bound to NRG1β[42] but with the equivalent epitope being

unoccupied. The fact that the receptor only exists in one state allows it to be poised for interacting with the other members without the need of any conformational change.

The **TKD** of ErbB receptors is much smaller in size than the ECD and consists of two lobes; the N-lobe which is constituted by β-sheets and the C-lobe which is mainly helical. The role of the TKD is to bind ATP, upon activation of the receptor by ligand binding, in the catalytic cleft located between the two lobes and to catalyze the transfer of the γ-phosphatase of ATP to tyrosine residues on the C-terminal tail which in turn will be served as a docking site for downstream effectors. Some features important for receptor activation are: the αC-helix on the N-lobe, which coordinates with a key residue (K721) within the cleft to stabilize ATP binding, the activation loop (A-loop) on the C-lobe and the conserved Asp-Phe-Gly (DFG) motif located on the N-terminus of the A-loop.

Questions regarding the activation mechanism of these receptors where first raised when Gotoh *et al.*[43] showed that phosphorylation of the A-loop of ErbB receptors is not required for catalytic activity. For the rest of the members of the RTK superfamily, however, the phosphorylation of the A-loop is essential for activation, as the receptors cannot be released from their *cis*-autoinhibited state without it[1]. It is worth noting that even though the mutation of Tyr845 (located on the A-loop) of EGFR to Phe did not abolish kinase activation[43], it has been found to be important for cell proliferation and transformation in breast cancer cells[44] as well as for EGF-induced DNA synthesis[45].

The same observation regarding the A-loop-independent activation of ErbB1 was also made later on when the first crystallographic structure of EGFR TKD was resolved by Stamos *et al.*[46]. Despite the fact that the TKD of EGFR was in an active-like conformation and structural features resembled those seen in other RTKs, the A-loop was not phosphorylated. The notion that the mechanism of ErbBs activation might be distinct from the other RTKs was further supported by yet another structure. In that structure, the TKI Lapatinib was co-crystallized with the EGFR TKD and revealed a novel conformation[47] which has been previously observed for the inactive state of both cyclin-dependent kinases (CDKs) and proto-oncogene tyrosine-protein kinases (Src)[48].

The aforementioned structures suggested that the TKD, as the ECD, exists in two different states as depicted in ***Figure 1-6***. In the inactive conformation (***Figure 1-6A***),

the αC-helix is displaced away (αC helix-out) from the ATP binding pocket by the small helical DFG motif. Furthermore, the flexible A-loop also points out in order to allow substrate (C-terminal tail) binding to the binding site which is located just outside of the cleft. In the active conformation (***Figure 1-6B***), the helical structure of the DFG motif no longer exists allowing αC-helix to be positioned in a close proximity with K721 and stabilize the active conformation.

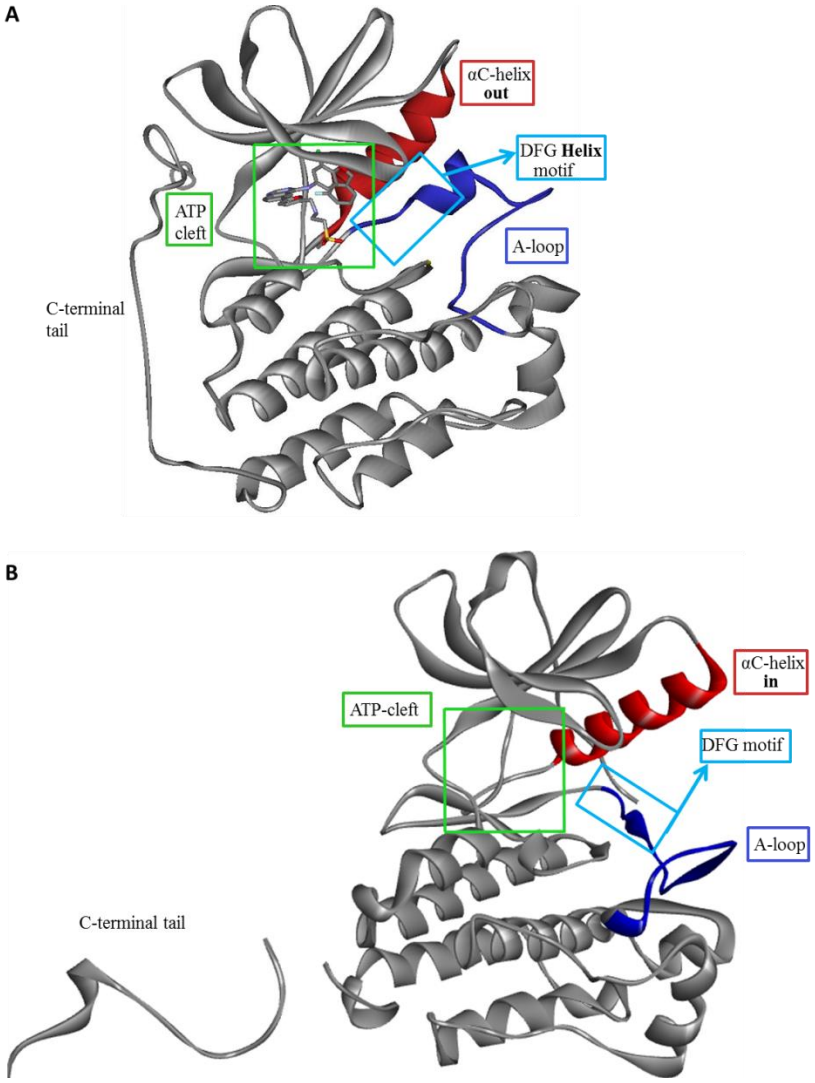


Figure 1-6: Crystal structures of the inactive and active conformation of the TKD of EGFR. (**A**) Crystal structures of wt EGFR bound to Lapatinib in the inactive conformation (PDB: 1XKK). (**B**) Crystal structure of wt EGFR in the active conformation (PDB: 1M14). In both structures, the ATP-site (green framerwork), the αC-helic (red), the DFG motif (light-blue framework) and the A-loop (blue) are denoted.

These findings, along with several other structural and biochemical studies which untangled the mechanistic aspects of the activation of EGFR family[49–53], suggested an allosteric mechanism previously described for CDKs[49].

Upon ligand binding to the ECD, dimerization is promoted which brings the kinase domains of each monomer in close proximity and results in conformational changes allowing the formation of an asymmetric dimer between them. This dimer presents a head-to-tail configuration, where the C-lobe of one kinase ("activator" or "donor") interacts with the N-lobe ("receiver" or "acceptor") of the other kinase (***Figure 1-7A***). These hydrophobic interactions between the TKDs cause disruption of the *cis*-autoinhibitoty interactions of the receiver (αC-helix in, extension of the A-loop and deformation of the helical structure of the DFG motif) which can now bind ATP and catalyse the transfer of the phosphate group. In this dimer, the "activator" molecule is in the inactive conformation as in ***Figure 1-6A*** whereas the "receiver" in the active (***Figure 1-6B***). Importantly, the crystal structure of the asymmetric dimer showed that the C-terminal tail of the activator was positioned at the active site of the receiver, indicating phosphorylation in *trans*[49].

In contrast to the ECD, the inactive TKD has been crystallized in a dimeric arrangement[49,54] (in addition to the monomeric form shown above). This symmetric back-to-back dimer comprises TKDs which are both in the inactive-like state as none of the interactions which are crucial for activation take place (***Figure 1-7B***) and does not contribute to the phosphorylation of the receptor[49]. It should be noted that in the structure reported by Jura *et al.*[54], the symmetric interface interaction is mediated by the small AP-2 helices (residues 969-978), formed at the beginning of the C-terminal tail of each monomer (***Figure 1-7B***), suggesting an auto-inhibitory role of the tail[1].

Despite the fact that structures of the active TKD conformation of EGFR[49], ErbB2[55] and ErbB4[51] have been resolved, ErbB3 has never been crystallized in that state. Perhaps the reason for that is that ErbB3 N-lobe (asymmetric dimer interface) residues are divergent from those of the other members[56] and this prevents its TKD from acting as receiver in the asymmetric dimer. However, since the phosphorylation of ErbB receptors occurs in *trans* and the TKD of ErbB3 can function as an activator (C-lobe sequence is very similar to other receptors) when it heterodimerizes, ErbB3 is found phosphorylated in cells. On the other hand, its inactive conformation[56] is the same as the equivalent structures seen for the other ErbB family members[47,51,57].

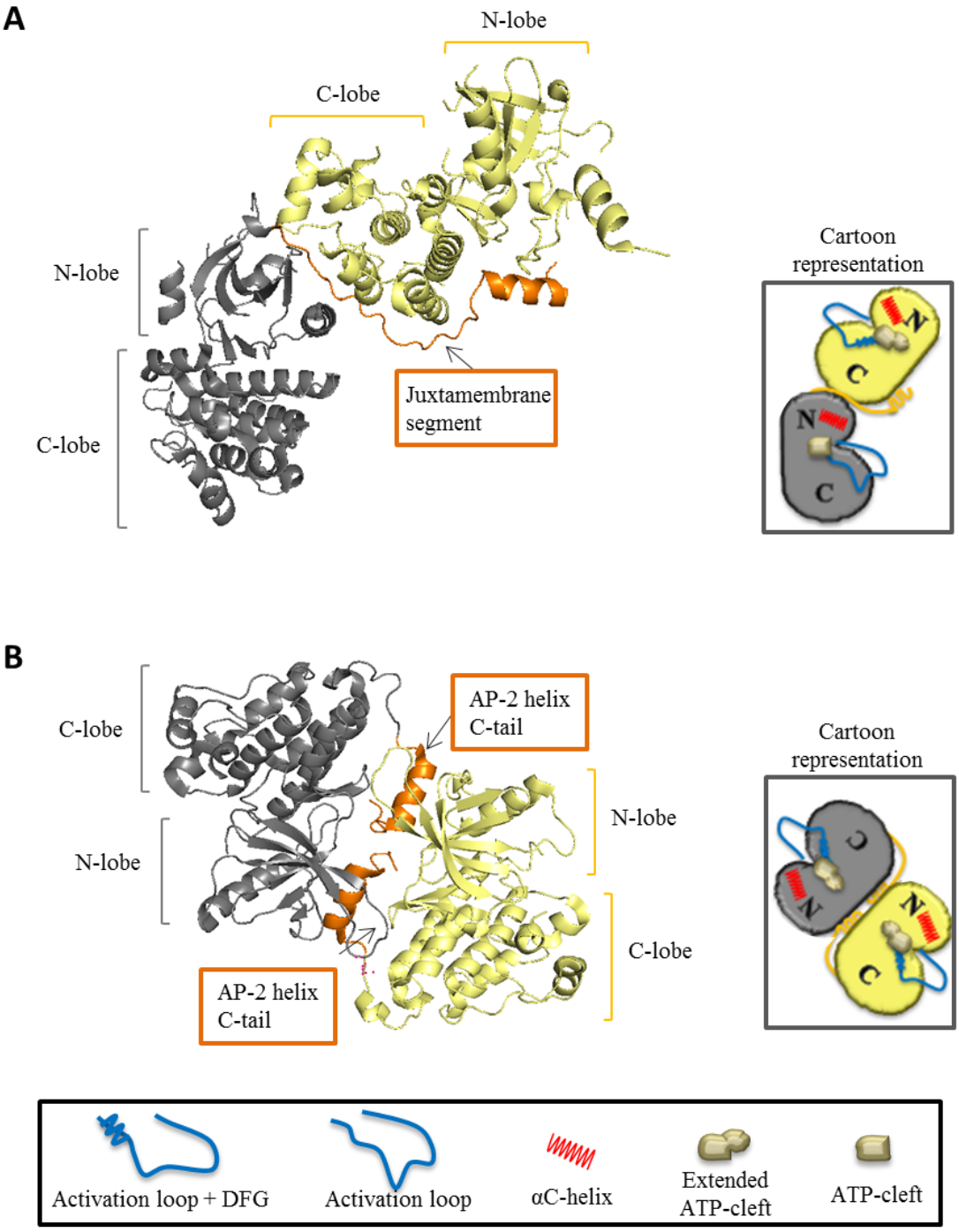


Figure 1-7: Crystal structures and cartoon representations of the asymmetric and symmetric TKD dimer. In the head-to-tail interaction of the asymmetric dimer (**A**, PDB: 3GOP), the JM (orange) latches the C-lobe of the activator (light yellow). In the symmetric dimer (**B**, PDB: 3GT8) the back-to-back interaction between the monomers is mediated by the small helices of their C-tail (orange).

Another structural feature that has been highlighted to be indispensable to the allosteric activation described above is the **JM**[53,54,58]. The JM can be separated in two regions: the JM-A (residues 645–657) which is helical and the more flexible JM-B

(residues 658–671). In the inactive state (monomers or symmetric dimers), JM-A has been predicted to be embedded in the inner leaflet of the membrane, possibly by interacting with membrane components such as Phosphatidylinositol 4,5-bisphosphate (PIP_2)[59–61], and to be monomeric[62,63]. Upon activation, ECD conformational rearrangements cause its release from the membrane and subsequently the formation of an anti-parallel helical dimer which contributes to the stabilization of the asymmetric dimer. It has also been demonstrated that different ligands induce distinct JM-A dimers[64], further underlying their importance in activation. On the contrary, JM-B is directly involved in the asymmetric dimer: the JM-B of the receiver cradles the C-lobe of the activator as it is shown in ***Figure 1-7A***. Mutations that disrupt the interaction between the JM-B and the C-lobe or deletion of the JM completely abrogate phosphorylation of EGF.

1.2.2 Activation Mechanism of ErbBs in Cells

Even though the organization of both inactive and active receptors, as well as their architecture on cells is yet to be defined, a simplified mechanism summarizing the aforementioned findings on the inactive-active transition of ErbBs is illustrated in ***Figure 1-8A***.

In the basal state, the receptors (apart from ErbB2) adopt the tethered conformation on the ECD with an inactive-like TKD conformation. While results from some studies question the existence of inactive pre-dimers[62,65], others have shown that ErbBs not only pre-dimerize even in the absence of ligand[66–68] but they may also form higher order oligimers[69–71]. Based on these contradicting results, the inactive receptor is represented in equilibrium between a tethered monomer and a tethered dimer with a symmetric TKD configuration. The presence of the tethered receptor in the inactive pre-dimers was recently demonstrated by Nevoltris *et al.*[72] using novel nanobodies which recognize the tethered conformation. It is also worth mentioning that the observed inactive dimers where heterodimers of EGFR (tethered) and ErbB2 which is an extended-only receptor.

In the activated state, ligand binding to monomers or pre-dimers leads to an extended 2:2 ligand-bound ECD dimer which promotes the formation of the asymmetric dimer on the intracellular and allows signal propagation. As for the basal receptor, oligomerization or clustering of the activated receptors has also been proposed[4,70,71,73].

The existence of higher order oligomers might also explain how ErbB2 can be phosphorylated in a ErbB2-ErbB3 system since the phosphorylation occurs in *trans* and ErbB3 is catalytically inactive[74].

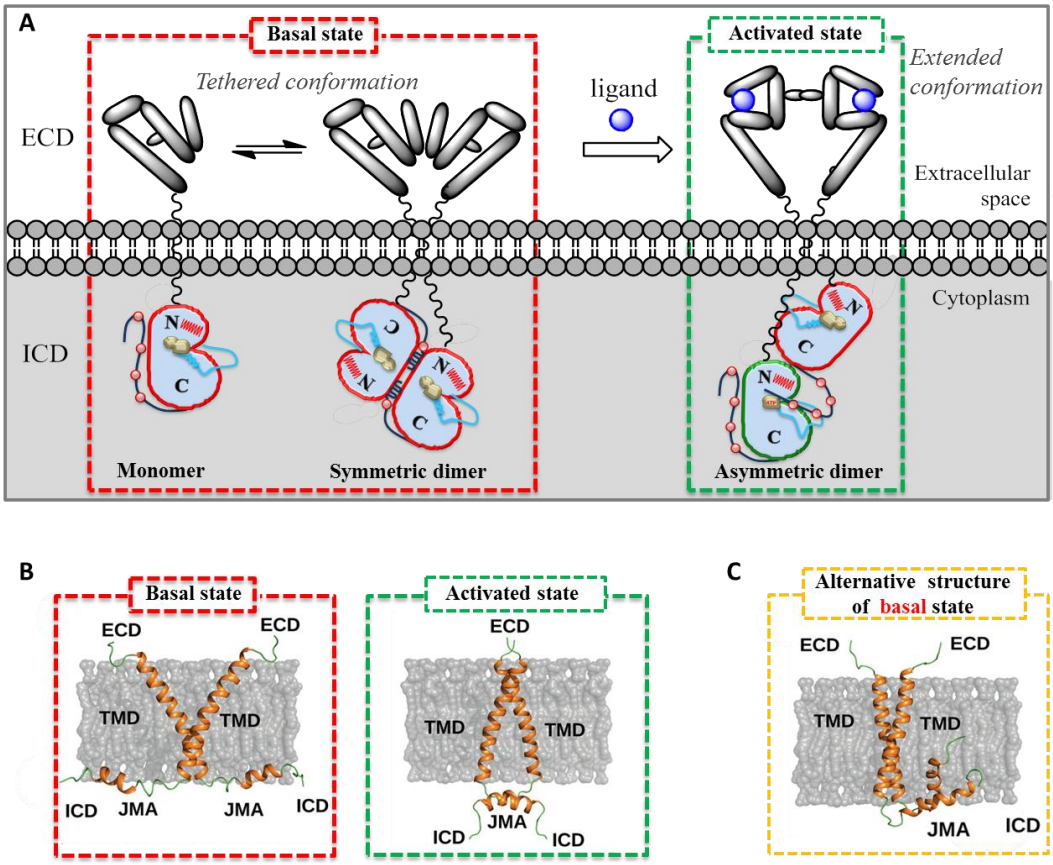


Figure 1-8: Mechanism of activation ErbB members and TM-JM dimerization crossings on cell surface. (**A**) In the basal state, receptors exist as either monomers or inactive pre-dimers. Upon ligand binding, ECD is released from the tethered conformation, dimerized and allow the asymmetric KD dimer formation. (**B)** Conformation of TM-JM in the basal (left: C-terminal crossing) and the activated (right: N-terminal crossing) state of dimeric receptor (Adapted from Bragin *et al*, *J Mol Biol*, 2016[75]). (**C)** Proposed alternative conformation of TM-JM in the basal dimeric receptor (Adapted from Bragin *et al*, *J Mol Biol*, 2016[75]).

The structural component responsible for the conformational coupling between the ECD and ICD is the **TM** domain. The TM is approximately 20 amino acid long with a helical structure and its association, in both homo- and hetero-dimers, has been suggested to contribute to dimer stabilization and activation[76–80]. All ErbB members have conserved motifs in both the N- and C-terminal of the TM (apart from ErbB3 which lacks the C-terminal motif) consisting of a GG4-like (GxxxG) sequence[15] which are involved in both asymmetric and symmetric dimer, respectively. In the companion articles by Endres *at al.*[62] and Arkhipov *et al.*[63], the authors employed *in silico*, Nuclear

Magnetic Resonance (NMR) and mutational studies to determine the TM-JM dimer conformation in both the unliganded-symmetric and extended-asymmetric dimers. According to this model, the TMs of the inactive dimers interact with each other at the C-terminus (C-crossing) of each monomer whereas in the active form it is a N-crossing (***Figure 1-8B***) conformation. In the former configuration the JMs are kept apart and embedded into the membrane while in the latter the N-crossing conformation causes release of JMs from the membrane and also favours the formation of the anti-parallel JM-A dimer.

A third, distinct, conformation of the TM-JM dimer in the inactive receptor was proposed earlier this year by Bragin *et al.*[75] (***Figure 1-8C***). In this structure, the crossing of TMs is slightly shifted, compared to C-crossing, towards the middle which results in a much smaller crossing angle. This TM crossing also affects the JM conformation which, in contrast to the previous model, can now dimerize in a parallel fashion with a part of it buried in the plasma membrane. The authors also state that this alternative inactive conformation is reminiscent of other receptors and that its transition to the active is relatively easier compared to the C-crossing.

Despite these fundamental discoveries in the last 15 years, the ErbB system still remains a complex network with many unknown aspects which need to be addressed for more accurate and efficient therapies. For instance, the role and the importance of the TM in the activation have been highlighted very recently by different studies[22]. Moreover, features such as higher-order oligomers, the role of the C-terminal tail and the conformation of the receptor in the different states are still unclear. Chemical biology approaches for developing novel probes along with fluorescent microscopy techniques comprise a promising combination to further explore such systems.

1.3 Fluorescence Microscopy

Fluorescence microscopy is a powerful imaging tool for modern biology and medicine. It is a variation of the conventional light microscopy and its principle relies on molecules, known as fluorophores, that when excited with light of appropriate wavelength re-emit light (fluorescence emission) at longer wavelength. This allows the fluorescence emission to be separated from the excitation and specific labelling of targets with fluorophore allow these to be distinguished within a sample (***Figure 1-9A***).

A fluorescent microscope requires: a light source for producing photons with energy capable of exciting a specific fluorophore, an objective lens for focusing the light as well as for collecting the emitted light, a detector for the emitted photons and a filter(s) for splitting the excitation and emission lights. There are three categories of fluorophores that are most usually used for visualization of a protein of interest as illustrated in ***Figure 1-9B***: fluorescent proteins which can be fused onto other proteins, small organic dyes and quantum dots for tagging biomolecules.

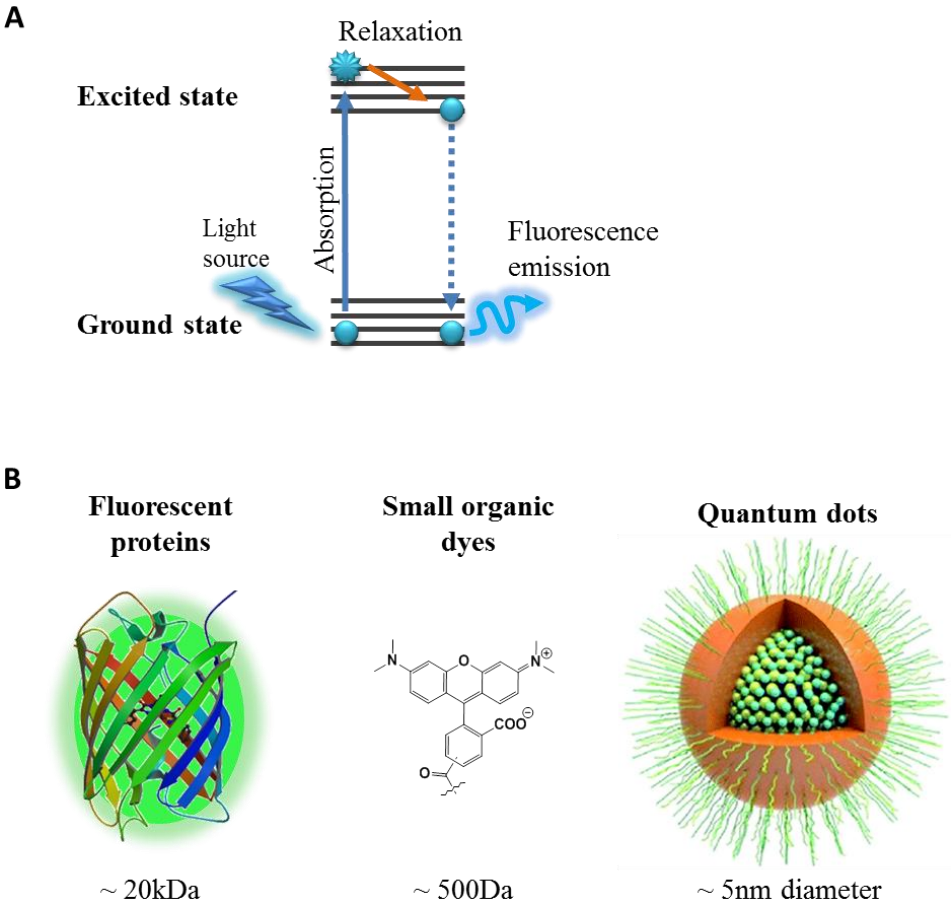


Figure 1-9: Jablonski energy diagram for fluorescence principle (A) and types of fluorophores (B).

Fluorescent microscopy techniques allow imaging of both fixed and live biological specimens and can provide insight on biological processes. These techniques are very sensitive to low numbers of labelled molecules, changes in the environment, *etc*. and enable scientists to conduct experiments under physiological conditions (e.g. temperature), directly detect a protein of interest (single-molecule imaging) in complex systems (e.g. cells) and extract structural and dynamic information (amongst other parameters)[81].

The two types of fluorescent microscopes used in the present study are scanning illumination (confocal) and surface illumination (total internal reflection fluorescence (TIRF)) which will be briefly described in the following two sections.

1.3.1 Confocal Laser Scanning Microscopy and Fluorescence Lifetime Imaging

The basic principles of confocal imaging were introduced by Marvin Minsky in 1955[82]. Although modern confocal microscopes have been highly evolved from the 'prototype', the principles they follow remain the same.

Confocal laser scanning microscopy (CLSM) is an advanced optical technique for visualization of a specimen with the use of fluorescent molecules and it has proved a powerful tool in cell biology as well as in other research fields. Confocal microscopy is based on the principle that out-of-focus glare can be eliminated or rejected by placing a pinhole in front of the detector (***Figure 1-10A***). The laser beam comes from the side, is filtered by the dichroic splitter in order to give the desired wavelength, reflected *via* the objective and excites the specimen. The emitted light from the specimen then returns back, is de-scanned and ends up through the pinhole to the detector. The construction of an image is made pixel by pixel and line by line as the laser scans the specimen. This pinhole is the core difference between confocal and conventional wide-field microscopy where the entire field is excited. This point-wise excitation and detection offers several advantages such as: rejection of out of focus fluorescence, scanning in x-y planes with different depths which allows imaging of thick samples without the need of physical sectioning and acquisition and reconstruction of 3-D images (***Figure 1-10B***).

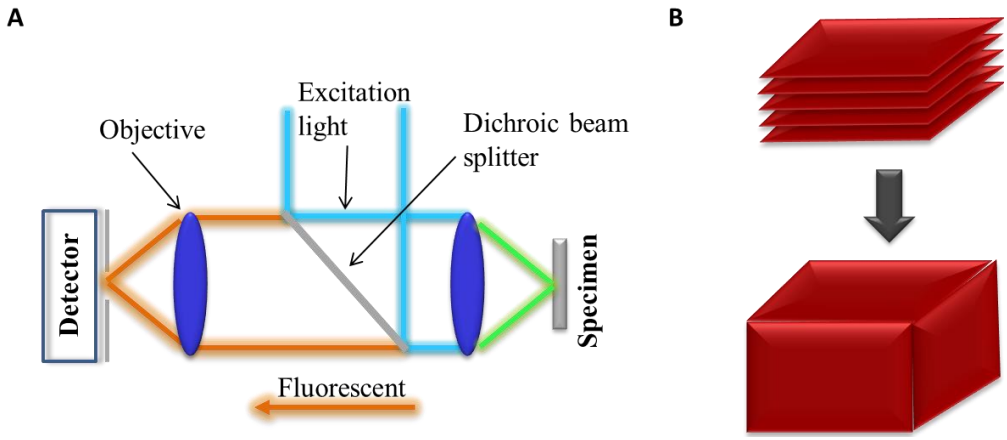


Figure 1-10: Principle of Confocal laser scanning microscopy. Illustration of the principles of (**A**) CLSM and (**B**) reconstruction of 3D image.

In biological sciences CLSM is used, but not limited to, for visualization of both live and fixed samples, localization and co-localization of proteins, quantitative analysis (e.g. binding studies), visualization of cell processes (cell division) and immunofluorescence. Many fluorescence imaging techniques are compatible with CLSM including: Fluorescence Recovery After Photobleaching (FRAP) for measuring diffusion of both cell membrane and interior components, Förster (or Fluorescence) Resonance Energy Transfer (FRET) for studying protein interactions, Fluorescence Lifetime Imaging (FLIM) which overcomes some limitation of intensity-based FRET and is sensitive to environment changes (e.g. pH) and Stimulated Emission Depletion (STED) a super-resolution technique which by-passes the diffraction limit of conventional microscopes[83].

FRET is a process whereby the energy from one excited fluorophore (donor) is transferred to another (acceptor). As it can be used as a biophysical measurement for determining molecular proximity within ~10 nm distance[83], FRET is most commonly used to study protein-protein (or -ligand) interaction and protein conformation[84]. The principle of FRET is shown in ***Figure 1-11***. When an excited donor fluorophore is in close proximity to an acceptor fluorophore, the energy of the donor before returning to the ground state can be transferred to the acceptor and excite it. The efficiency of the FRET process between an isolated donor and acceptor depends upon the donor quantum yield, relative orientation of the donor and acceptor dipoles, the overlap between the donor emission and the acceptor spectra and the distance between and acceptor.

FRET can be quantified by measuring the reduction in the donor fluorescence intensity and sensitised acceptor emission, depolarization of fluorescence emission or by the apparent reduction of the fluorescence lifetime of the donor[85]. The fluorescent lifetime (τ) of a fluorophore is the mean time a molecule stays in the excited state before returning to the ground state. Measurement of fluorescence lifetimes can be done either in the frequency domain or the time domain[86]. The frequency domain approach is compatible with widefield microscopy methods and the time domain is compatible with point scanning microscopy. Determining the fluorescence lifetime at every pixel of a fluorescence image is known as Fluorescence Lifetime Imaging Microscopy (FLIM).

The advantages of using FLIM to quantify FRET (FLIM-FRET) over intensity-based FRET are that only the donor is monitored[87] and that artefacts from signal cross-contamination, concentration and photobleaching are greatly reduced[83,88].

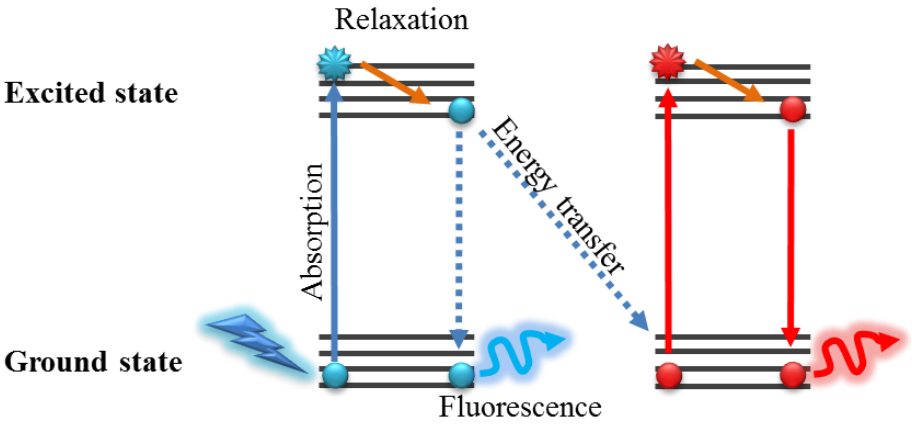


Figure 1-11: Förster resonance energy transfer principle.

1.3.2 Single Molecule Imaging using Total Internal Reflection Fluorescence (TIRF) Microscopy

Imaging of single molecules was first reported nearly 30 years ago by two independent groups who visualized solid crystals[89,90]. However, the need for aqueous media for single-molecule imaging of biological samples set a barrier to this method[90]. To overcome this problem, Funatsu *et al.*[91] exploited TIRF microscopy which had been earlier introduced by Axelrod *et al*[92].

TIRF microscopy is widely used for imaging the membranes cultured cells and thin samples. TIRF can be operated on a wide-field epi-fluorescence microscope with the use of an objective lens with a high Numerical Aperture (NA). The laser beam is focused onto the back aperture of the objective lens so that it emerges from the lens as a collimated beam. By translating the position of the beam away from the centre of the back aperture, the beam will emerge from the objective at an angle. Once the angle is large enough the beam will be totally internally reflected (TIR) at an interface of media with different refractive indexes e.g. glass-sample medium (***Figure 1-12***). The critical angle for TIR depends upon the refractive index mismatch. The sample is then illuminated by the evanescent wave generated by TIR that penetrates the sample to a depth of only a few hundred nanometers. This results in a large reduction of the background signal.

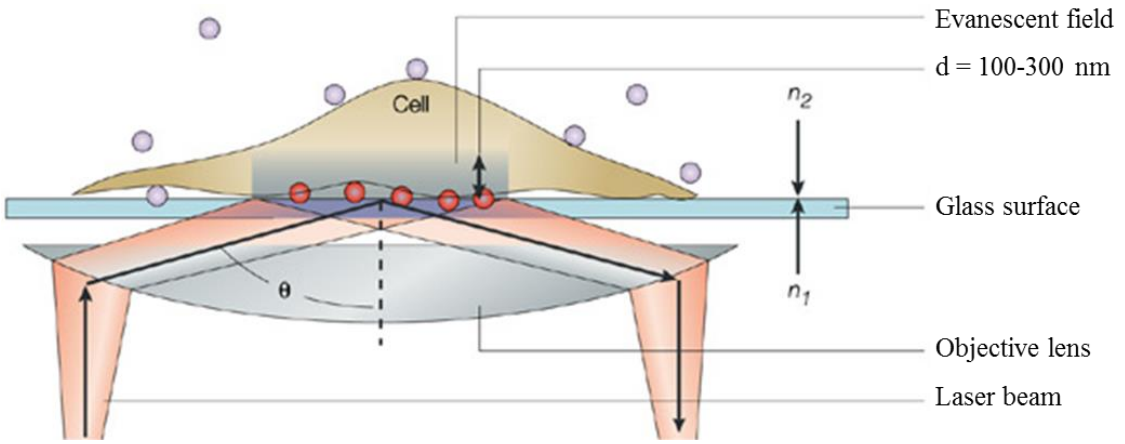


Figure 1-12: Schematic representation of Total internal reflection fluorescence (TIRF) microscopy (Adapted from Y. Sako and T. Yanagida, Nat. Rev. Mol. Cell Biol., 2003[93]).

Single-molecule methods, either using confocal[45] or TIRF microscopy[94], allow one to visualize, detect and analyse individual labelled particles. This feature of SMMs enables scientists not only to extract information regarding the mean behaviour of a certain property of the system under study but to examine the heterogeneity of a system as even rare events are reported[95]. Furthermore, the development of single-molecule localization algorithms[70,96] and the use of Charge-Coupled Device (CCD) cameras, allows the localization of molecules with a precision below the diffraction limit of light and the detection of molecule movement, or other events, with millisecond time resolution. The most commonly used single molecule methods compatible with TIRF are: single-particle tracking (SPT) for studying protein dynamics, localization-based methods for extracting information regarding the geometry and stoichiometry of protein complexes and single-molecule FRET (smFRET) for studying intra- and inter-molecular dynamics and kinetics[97].

SPT is a live-cell imaging method for studying protein dynamics by tracking the lateral movement of molecules and protein-protein interactions when 2- or 3-colour experiments are employed. Parameters that can be extracted from SPT are the diffusion coefficients of the particles, the confinement of the movement and the frequency and duration of colocalization events[95,98]. The main advantage of SPT over other methods for studying protein dynamics (such as fluorescence recovery after photobleaching (FRAP)[99] or fluorescence correlation spectroscopy (FCS)[100]) is, as mentioned above, the detection of events with spatial and temporal resolution[95]. SPT method has also been employed in the EGFR field and has provide insight into EGF-driven dimerization[101]

and pre-dimerization[102], activation mechanism of EGFR cancer mutant[103], to name a few.

Single-molecule localization methods on the other hand take place in either immobilized, on a glass surface, molecules or in fixed cells and provide information regarding the stoichiometry of molecules. The method that is employed herein is the Fluorophore Localization Imaging with Photobleaching (**FLImP**) method which has been developed by Martin-Fernandez's group[70,104] for determination of oligomer size and distribution. This super-resolution technique requires fixed samples and allows the measurement of the distribution of the separations between pairs of molecules with ~6 nm resolution. The advantage of FLImP over other super-resolution techniques such as photoactivated localisation microscopy (PALM)[105], stochastic optical reconstruction microscopy (STORM)[106] and stimulated emission depletion (STED)[107] is that it allows to measure inter-molecular distances in the range of ~10-60 nm as it has been demonstrated in previous studies[70,71].

1.4 Labelling Techniques for Intracellular Targets

The use of fluorescent microscopy techniques for studying biological systems requires specific labelling of the protein(s) of interest. Moreover, the labelling must not perturb the conformation of the protein, should ideally be compatible with live cell imaging and, ideally, be sensitive enough to monitor conformational changes. Even though the aforementioned requirements for labelling an extracellular target are met by many methods or probes, the same is not true for intracellular targets.

The study of the conformation of ErbBs kinase domains in a cellular context with imaging techniques is significantly hindered by the lack of suitable probes. The most commonly used labelling methods are illustrated in ***Figure 1-13***.

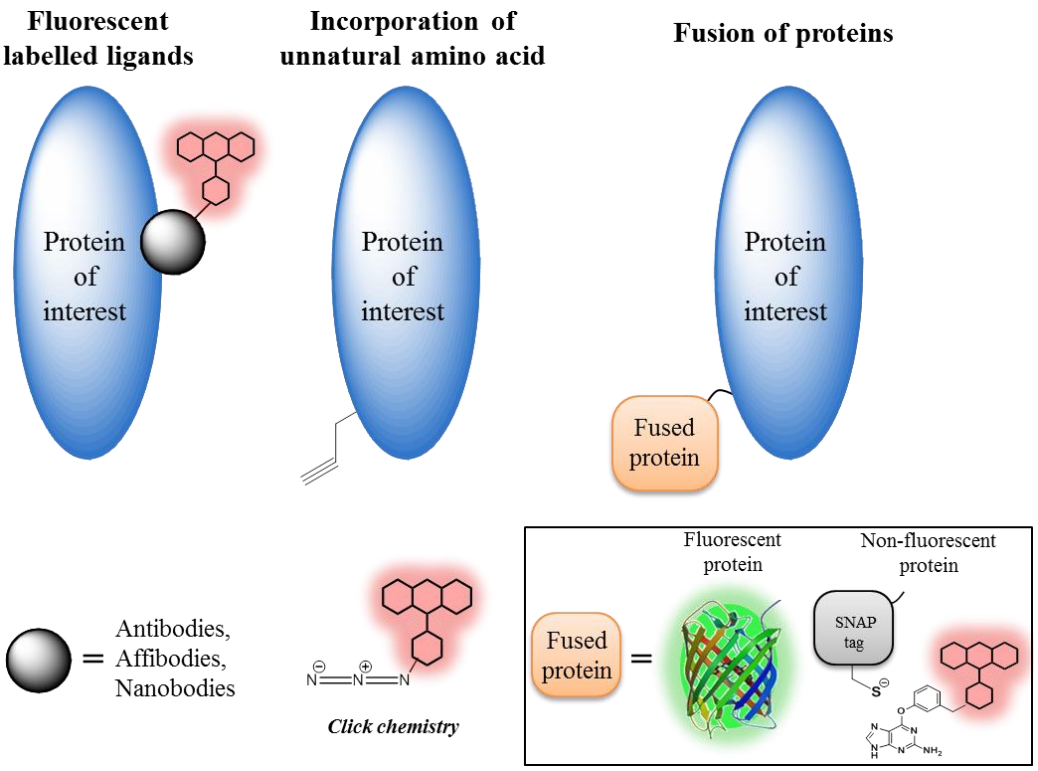


Figure 1-13: Most commonly applied labelling techniques. A protein of interest can be tagged by fluorescently labelled ligands (left) or genetically modified to be expressed with either unnatural amino acid for click chemistry (middle) or fused protein (right).

Antibodies (IgG) are proteins which are used in many imaging applications due to their high affinity (low nanomolar or even at the picomolar range[108]) and specificity but also because they are widely available commercially. However, an intact IgG protein has a molecular weight of ~150 kDa (similar to full-length EGFR) and is further modified by glycosylation on different sites, a modification which is essential for its functions[109]. Due to their size, they are unable to pass the cell membrane and therefore labelling of intracellular proteins is only possible after cell fixation and permeabilization, which limits their use. Besides, conformation-sensitive antibodies are hard to produce and select, even though they have been successfully developed for some targets[110]. Intracellular expression of non-antibody high affinity binders such as nanobodies and monobodies, which are much smaller in size, has been performed in some studies[111–113]. However, their applications in imaging require fusion to fluorescent proteins (FP) which increases their size. Moreover these reagents still need to be produced and selected on-demand for each new protein target.

Incorporation of unnatural amino acids in the sequence of a protein is an emerging method for specific labelling[114–116]. The side chains of the unnatural amino acids, which can be introduced by expanding the genetic code[117], bear a functional group capable of participating in exceptionally specific reactions with high kinetics rates (bioorthogonal chemistry) *in vitro*, under mild conditions[118]. This approach, however, requires the introduction of two plasmids (mutant protein and modified tRNA) and the supplementation of the culture with an unnatural amino acid bearing the reactive group to ensure the insertion of the tag in the desired position. While this approach can guarantee high labelling specificity, insertion sites need to be carefully planned in order to avoid disrupting important interaction interfaces. Another more critical challenge is the low efficiency of the incorporation of unnatural amino acids[119] and hence the labelling efficiency. Moreover, this technique is only suitable for the study of ectopically expressed kinases and therefore not ideal for the study of primary samples.

The same caveats apply to the production of fusion constructs labelled with fluorescent proteins or other tags (based on enzyme-labelled substrate interaction), with the additional concern that the size and position of the tag might compromise the function of the fusion protein[98,120,121].

1.4.1 Site-specific Labelling Using Small-molecule Probes

The limitations of the available labelling techniques for intracellular targets have led scientists to explore alternative methods. Such a method is site-specific labelling by small-molecule probes.

Small-molecule imaging probes are low-molecular weight compounds which can be very specific to their primary target protein and allow the visualization of the protein of interest. The development of such probes was inspired by the emerging field of activity-based protein profiling (ABPP)[122], pioneered by Benjamin Cravatt[123]. Probes for both applications (imaging and ABPP) share the same principle in terms of their design and structure; they consist of a reactive warhead, a linker and a tag (***Figure 1-14A, top***)[124–126]. The reactive warhead is responsible for the recognition of the target whereas the choice of the tag (a dye or an affinity label) depends on the application (labelling or chemical proteomics respectively). The role of the linker (PEG, amino acid chain or cleavable linkers) is to keep apart the tag from the warhead so they do not interfere with

each other, as such an interference would likely alter the specificity of the probe. Such probes can be either synthesized with tag attached (***Figure 1-14A, top***) [124–126] or prepared with a functional end (e.g. clickable linker) on a, usually, shorter linker which can be tagged *in situ* (bioorthogonal reaction, ***Figure 1-14A, bottom***)[127,128]. An advantage of the latter (bioorthogonal probes) is that they can be used for both labelling and affinity-based pull-down experiments as the choice of the tag takes place after binding of the probe. Bioorthogonal probes have been also used in order to circumvent the altered specificity-affinity and/or permeability of pre-conjugated probes[129].

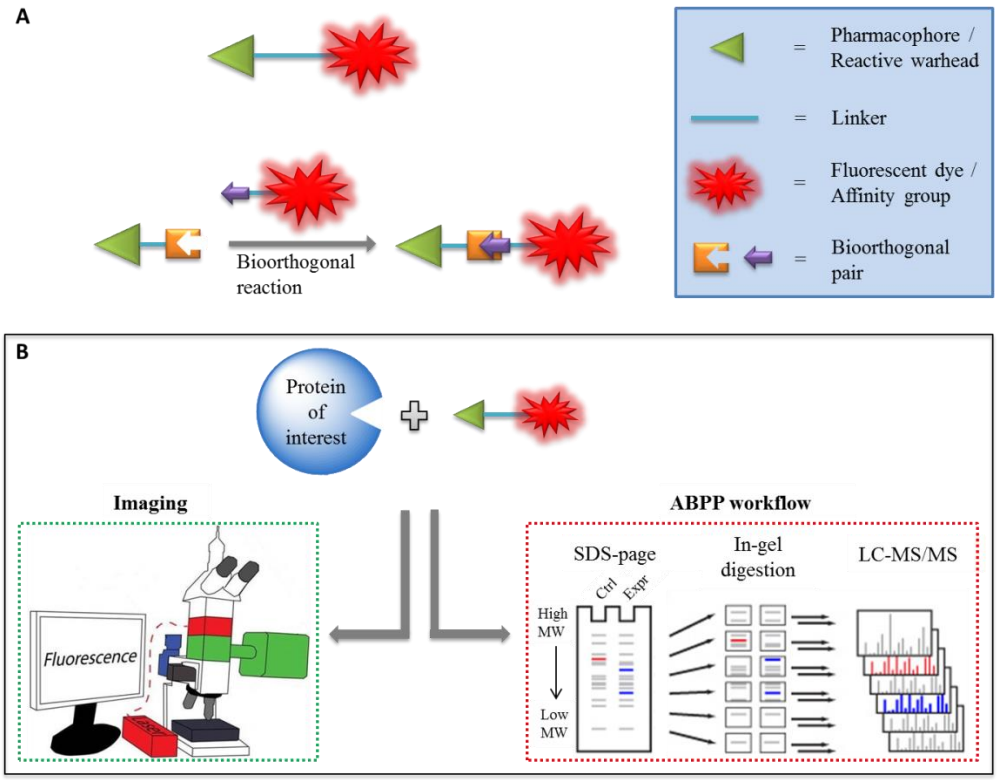


Figure 1-14: General structure of small-molecule probes and their applications. (**A**) Small-molecule probe consists of a warhead which allows specific binding in a target, a linker and a fluorescent dye or an affinity group. These probes can be either synthesized with the tag attached (top) or prepared with a functional group which can be participated in bioorthogonal reactions (bottom). (**B**) Their main application after labelling of the target are visualization using Imaging techniques (left) or in Gel analysis (right).

The two main techniques where such probes have been used are ABPP and imaging (***Figure 1-14B***). Applications of both methods offer information in a variety of studies such as: inhibitor target identification[130], binding studies[131], off-target profiling[127], kinetic studies[132], drug distribution[126], to name a few. Apart from *in vitro* experiments, the imaging approach has also been applied for *in vivo* studies in mice models[133–135].

This alternative labelling approach for studying biological processes with fluorescent microscopy seems to be promising, especially due to its suitability for labelling intracellular targets. Among the few examples in literature reporting the development of such probes for imaging, only a small number of them have reported novel biological information regarding their target protein, while most concerned the characterization of the recognition element of the probe. Therefore, this new method may offer new opportunities, as well as challenges, in the field of fluorescent microscopy.

2 Aim of Thesis

The main objectives of this project were: 1) the desing and synthesis of fluorescent TKIs (fTKIs) for ErbB family members (***Chapter 4***), 2) their evaluation and the development of methods for the use of fTKIs in fluorescence microscopy (***Chapter 5***) and 3) their applications in order to get a better insight into the drugs mechanism of action in cell-based assays and the basal EGFR architecture on cells (***Chapter 6 & 7***) (***Figure 2-1***).

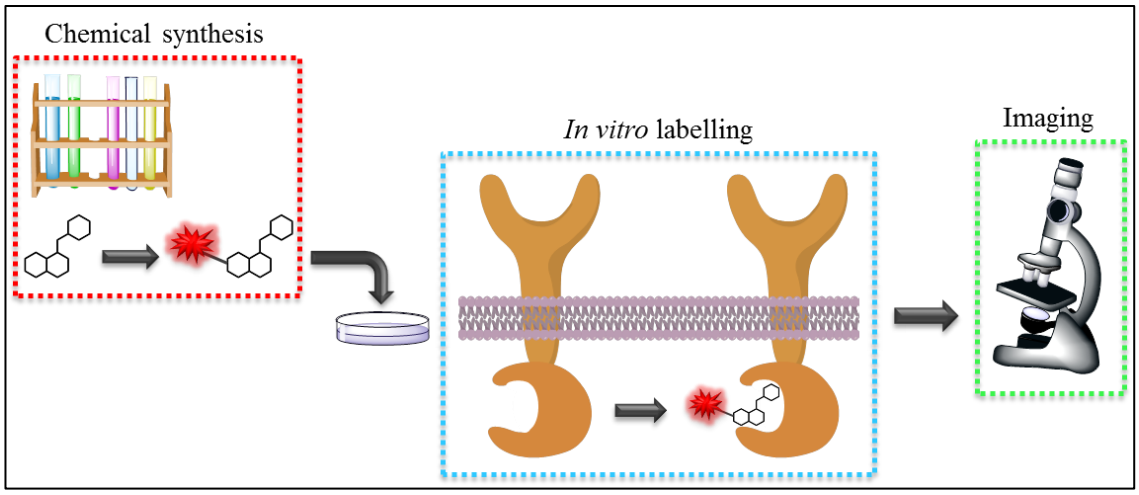


Figure 2-1: Schematic representation of project's aim.

To achieve the first objective, I have performed a literature review on ErbBs' TKIs in order to identify the most potent and selective compounds as well as to help me rationally design the synthesis of the labelled compounds. The modified compounds were synthesized based on the reported synthesis of the parent compounds. In terms of the second objective, I have evaluated the suitability of the synthesized fTKIs for use as imaging probes (binding affinity, cell permeability, specificity, *etc*) and have developed and validated methods for their use in different fluorescence microscopy techniques. Finally, for the third objective I have exploited one of the developed methods and combined it with methods for probing the EGFR ECD structure (previously reported by my group[70,71,136,137]) in order to investigate both drugs sensitivity/resistance mechanism in cell-based assays and the basal EGFR architecture on cell surface.

3 Materials and Methods

3.1 Chemistry

3.1.1 General experimental

All reagents, chemicals and solvents (HPLC grade) were purchased from Sigma-Aldrich, Alfa Aesar, Fisher Scientific, Fluka or Invitrogen and used without any further purification.

Standard syringe techniques were applied for transfer of air/moisture sensitive reagents and dry solvents. Reaction temperatures were monitored and controlled with an electronic contact thermometer on the outside wall of the reaction vessels. Reaction times refer to hold times at the indicated temperatures.

Chromatographic purification was carried out on Biotage Isolera flash chromatography using normal and reverse-phase Biotage SNAP cartidges. Thin-layer chromatography (TLC) analysis was performed with Kieselgel 60 F254 (Merck) plates and visualized using UV light.

^{1}H and ^{13}C NMR spectra were recorded on Bruker Advance DPX 2500 or DRX 500 instrument using deuterated solvents. Chemical shifts (δ) on ^{1}H spectra are referred to tetramethylsilane (TMS) and given in parts per million (ppm). Coupling constants are referred given in Hz. DEPT, COSY and HMQC measurements were also used to aid the peak assignment.

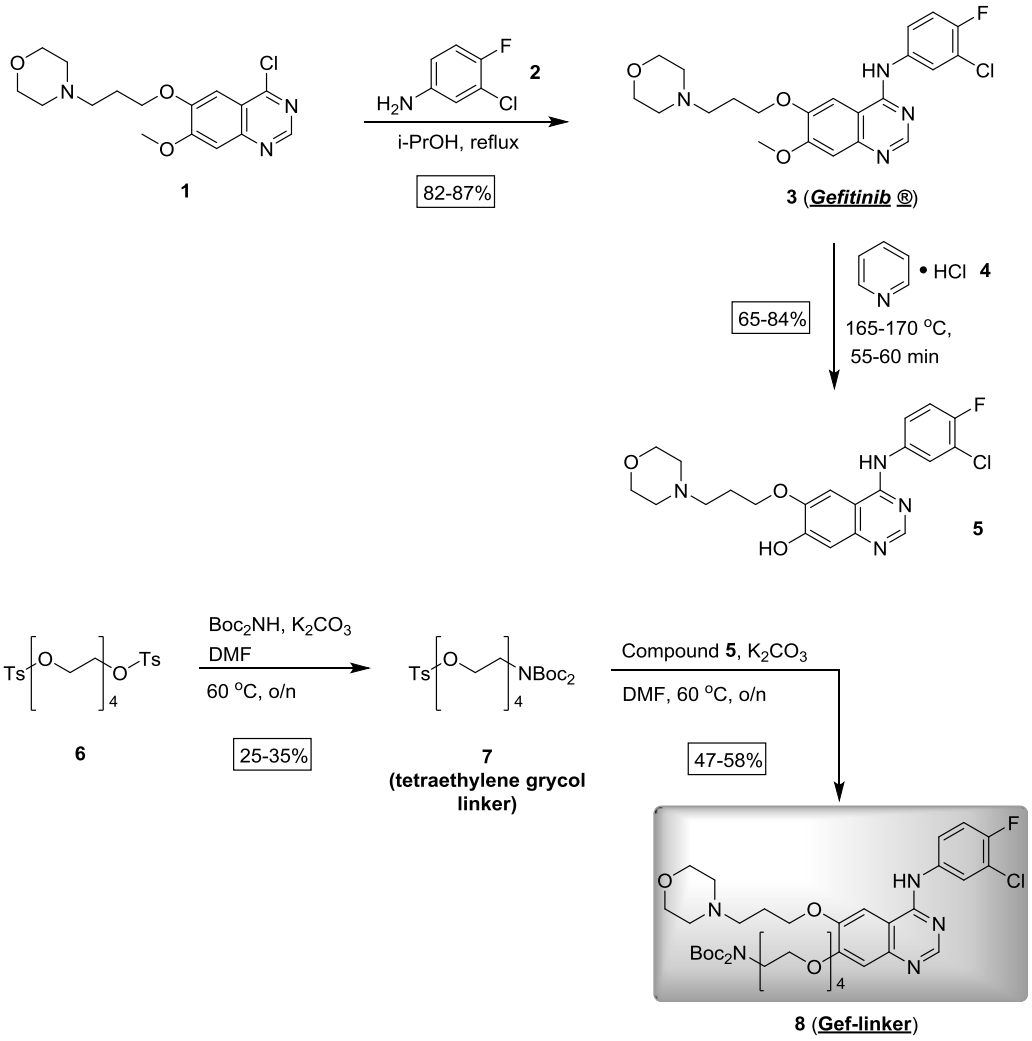
Analytical HPLC-MS was performed on Shimadzu LCMS-2010EV systems using reverse phase Atlantis dC18 columns (3 μm, 2.1 x 50 mm), gradient 5-100% B (A = water/ 0.1% formic acid, B = acetonitrile/ 0.1% formic acid) over 3 min, injection volume 3 μL, flow = 1.0 mL/min. UV spectra were recorded at 215 nm using a Waters 2788 dual wavelength UV detector. Mass spectra were obtained over the range m/z 150 to 850 or 150 to 1300 at a sampling rate of 2 scans per second using Waters LCT or analytical HPLC-MS on Shimadzu LCMS-2010EV systems using reverse phase Water Atlantis dC18 columns (3μm, 2.1 X 100 mm), gradient 5-100% B (A = water/ 0.1% formic acid, B = acetonitrile/ 0.1% formic acid) over 7 min, injection volume 3 μL, flow

= 0.6 mL/min. UV spectra were recorded at 215 nm using a Waters 2996 photo diode array.

High resolution mass spectrometry (HRMS) was performed at the Mass Spectrometry Facility (King's College London). Samples were flow injected in positive ESI mode. The mobile phase used was A: 0.1% formic acid in water (HPLC grade) and A: 0.1% formic acid in MeCN (HPLC grade) with flow rate 200ul/min.

3.1.2 Compounds synthesis

3.1.2.1 Synthesis of Gef-linker



Scheme 3-1: Total synthesis of Gef-linker (8) compound.

Compound 3 (Gefitinib)[138]**:** In a sealed tube containing 4-Chloro-6-[3-(4-morpholinyl)propoxy]-4-quinazoline (**1**) (1.0 g, 2.97 mmol) in *i*-PrOH (15 ml), 3-chloro-4-fluoro-aniline (**2**) (645 mg, 4.46 mmol) was added. The mixture was refluxed until precipitate was formed (approx. 3.5 h) and then cooled to an iced bath. The precipitate was collected by filtration, washed with chilled *i*-PrOH and dried over air to afford a white powder. The collected filtrate was evaporated *in vacuo* up to 10 ml and the above procedure was repeated 2-3 times. The filter cakes were collected together (1.0 g, 2.97 mmol, 85%) and the product was used in the next step without further purification (>95% purity). The ^{1}H NMR data for the compound (**3**) are identical to literature data[139]. MS (C18, ES$^+$) 447.1 $[M + H]^+$, $C_{22}H_{24}ClFN_4O_3$ requires *M* 446.1521; ^{1}H NMR (500 MHz, DMSO) δ 9.90 (s, 1H), 8.55 (s, 1H), 8.20 (dd, *J* = 6.7, 2.3 Hz, 1H), 8.06 (s, 1H), 7.92 – 7.86 (m, 1H), 7.46 (t, *J* = 9.1 Hz, 1H), 7.25 (s, 1H) 4.24 (t, *J* = 5.8 Hz, 2H), 4.02 (s, 3H), 3.81 (t, *J* = 4.41 Hz, 4H), 2.46 (t, *J* = 7.1 Hz, 2H), 2.39 (s, 4H), 1.99 (p, J = 6.7 Hz, 2H).

Compound 5: *Method I*[140]. To a sealed tube containing Gefitinib (0.4 g, 0.89 mmol) in anhydrous DMF (5 ml) lithium chloride (188 mg, 4.47 mmol) was added and the mixture was refluxed overnight. Then the mixture was cooled to room temperature and water (10 mL) was added. The mixture was extracted by EtOAc (3 × 10 ml), washed with brine (saturated solution of NaCl in water), dried over $MgSO_4$, filtered and concentrated *in vacuo* to give the crude product. Purification by column chromatography (10-40% EtOAc in heptane) gave the compound **5** as a white powder (42.4 mg, 0.10 mmol, 11%).

Method II[141]. The above procedure was repeated but the reaction was performed in a microwave oven for 9 h. The yield of the reaction was 42%.

Method III[142]. A sealed tube containing Gefitinib (0.5 g, 1.12 mmol) and an excess of pyridinium hydrochloride (**4**) was heated to 165°C for 55-60 min. Then water was added (8 ml) carefully to the hot mixture and left to cool at room temperature for 1.5 h. The precipitate was collected by filtration, washed with water and dried to afford a dark-grey solid (360 mg, 0.83 mmol, 74%) which was used to the next step without further purification (>95% purity). The ^{1}H NMR data for the compound (**5**) are identical to literature data[140]. MS (C18, ES$^+$) 433.1 $[M + H]^+$, $C_{21}H_{22}ClFN_4O_3$ requires

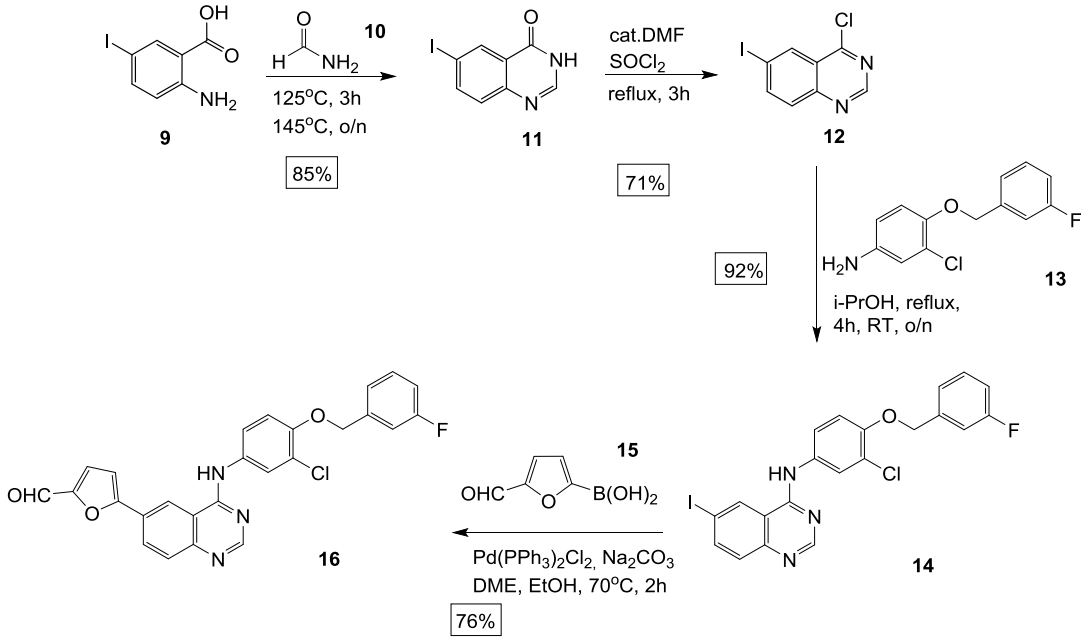
M 432.8758; ^{1}H NMR (500 MHz, DMSO) δ 8.45 (s, 1H, Ar), 8.02 (dd, *J* = 6.7, 2.6 Hz, 1H, Ar), 8.06 (s, 1H, Ar), 7.73 – 7.65 (m, 2H, Ar), 7.27 (t, *J* = 9.0 Hz, 1H, Ar), 7.17 (s, 1H, Ar), 4.26 (t, *J* = 5.8 Hz, 2H), 3.58 (m, 4H), 2.69 – 2.64 (m, *J* = 15.0 Hz, 2H), 2.57 (s, 4H), 2.16 – 2.09 (m, 2H).

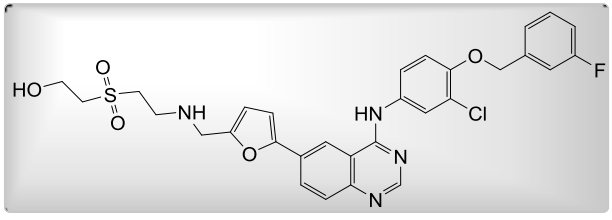
Compound 7 (tetraethylene grycol linker)[143]**:** To a solution of tetraethylene glycol di-*p*-tosylate **6** (3.50 g, 6.97 mmol) in anhydrous DMF (4 ml) di-tert-butyl iminodicarboxylate (**Boc$_2$NH**, 1.97 g, 9.06 mmol) and potassium carbonate (0.96 g, 6.97 mmol) were added. The resulting mixture was stirred at 60 °C for 4 h and then left to cool to room temperature, followed by addition of EtOAc (4 ml) and water (4 ml). The mixture was transferred to a separating funnel and the organic layer was washed with 1N HCl, water and brine. The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo* to give the crude product. Purification by column chromatography (10-30% EtOAc in heptane) gave the compound **7** as colorless oil (0.92 g, 1.67 mmol, 24%). The ^{1}H NMR data are identical to data reported previously by Shreder *et al*[143]. MS (C18, ES$^+$) 548.1 $[M + H]^+$, $C_{25}H_{41}NO_{10}S$ requires *M* 547.2451; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.72 (d, *J* = 8.3 Hz, 2H), 7.27 (d, *J* = 8.1 Hz, 2H), 4.08 (dd, *J* = 9.1, 4.1 Hz, 2H), 3.71 (dd, *J* = 11.5, 5.2 Hz, 2H), 3.63 – 3.59 (m, 2H), 3.55 – 3.46 (m, 11H), 2.37 (s, 3H), 1.42 (s, 18H).

Compound 8 (Gef-linker)[143]**:** To a solution of compound **5** (100 mg, 0.23 mmol) in anhydrous DMF (8 ml) compound **7** (253 mg, 0.46 mmol) and potassium carbonate (128 mg, 0.92 mmol) were added. The resulting mixture was stirred at 60 °C overnight and then left to cool to room temperature, followed by addition of EtOAc (4 ml) and water (4 ml). The mixture was transferred to a separating funnel and the organic layer was washed with 1N HCl, water and brine. The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo* to give the crude product. Purification by column chromatography (20-80% EtOAc in heptane) gave the compound **8** as yellowish oil (87 mg, 0.11 mmol, 47%). R_f = 0.61 (10% MeOH in DCM).; ^{1}H NMR (500 MHz, MeOD) δ 8.45 (s, 1H), 8.02 (dd, *J* = 6.7, 2.6 Hz, 1H), 7.72 (s, 1H), 7.69 (ddd, *J* = 8.9, 4.0, 2.7 Hz, 1H), 7.27 (t, *J* = 9.0 Hz, 1H), 7.17 (s, 1H), 4.36 – 4.29 (m, 2H), 4.26 (t, *J* = 6.1 Hz, 2H), 3.98 – 3.94 (m, 2H), 3.76 (ddd, *J* = 16.5, 8.4, 4.6 Hz, 8H), 3.68 (dd, *J* = 5.6, 3.8 Hz,

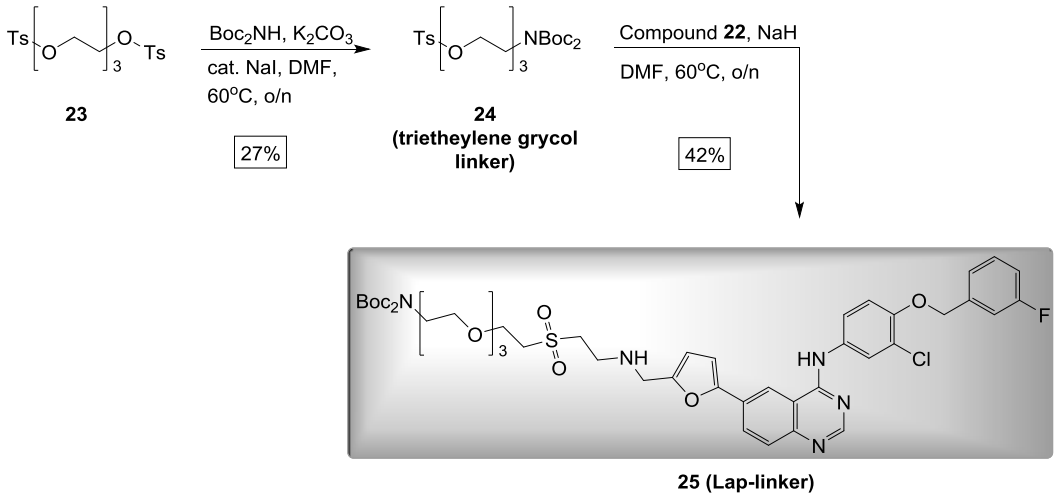
2H), 3.64 – 3.60 (m, 2H), 3.61 – 3.55 (m, 4H), 2.69 – 2.65 (m, 2H), 2.57 (s, 4H), 2.16 – 2.09 (m, 2H), 1.49 (s, 18H).; ^{13}C NMR (500 MHz, MeOD) δ 156.91, 155.65, 154.71, 153.22, 152.69, 152.41, 149.17, 146.10, 136.22, 136.19, 124.26, 122.38, 122.32, 119.95, 119.76, 116.08, 115.86, 109.16, 106.81, 102.59, 82.34, 70.60, 70.35, 70.27, 70.03, 68.52, 67.09, 65.87, 55.23, 53.19, 45.15, 26.91, 25.37.; HRMS (ESI) 808.3678 $[M + H]^+$, $C_{39}H_{55}ClFN_5O_{10}$ requires *M* 807.3621;

3.1.2.2 Synthesis of Lap-linker





22 (Modified Lap)



Scheme 3-2: Total synthesis of Lap-linker (25) compound.

Compound 11[144]**:** A solution of compound **9** (10 g, 38.02 mmol) in 80 ml of formamide **10** was heated under reflux for 3 h at 120-125°C and then at 145°C overnight. Upon cooling, the product was precipitated and collected by filtration. Finally, the crude solid was recrystallized from EtOH to give 8.95 g of compound **11** as light brown solid (yield, 85%). The ^{1}H NMR data of compound **11** are identical to literature data[145]. MS (C18, ES^{+}) 272.3 $[M + H]^{+}$, $C_8H_5IN_2O$ requires *M* 271.9447; ^{1}H NMR (500 MHz, DMSO) δ 1H NMR (500 MHz, DMSO) δ 12.40 (s, 1H), 8.39 (d, *J* = 2.1 Hz, 1H), 8.13 (s, 1H), 8.10 (dd, *J* = 8.5, 2.1 Hz, 1H), 7.46 (d, *J* = 8.5 Hz, 1H).

Compound 12[144]**:** To a flask containing compound **11** (8.95 g, 32.9 mmol) thionyl chloride (**$SOCl_2$**, 200 ml) was added dropwise at 0 °C. Catalytic amount of DMF (5-6 drops) was also added and the resulting mixture was heated under reflux for 3 hrs. Thionyl chloride was then removed under reduced pressure and the resulting residue was washed by slow addition of saturated Na_2CO_3. The product was extracted with EtOAc (3 × 100 mL) and the organic layers were dried over Na_2SO_4 and concentrated *in vacuo* to give the crude solid, yellowish needles, which was used to the next step without further purification (6.82 g, 23.5 mmol, 71%). The ^{1}H NMR data of compound **12** are identical to literature data[146]. MS (C18, ES^{+}) 290.7 $[M + H]^{+}$, $C_8H_5ClIN_2$ requires *M* 289.9108; ^{1}H NMR (500 MHz, DMSO) δ 1H NMR (500 MHz, DMSO) δ 9.07 (d, *J* = 5.3 Hz, 1H), 8.66 (d, *J* = 1.8 Hz, 1H), 8.21 (dd, *J* = 8.8, 1.9 Hz, 1H), 7.80 (d, *J* = 8.8 Hz, 1H).

Compound 14[147]**:** In a flask containing compound **12** (6.8 g, 23.4 mmol) in *i*-PrOH (160 ml), compound **13** (5.8 g, 4.46 mmol) was added. The mixture was refluxed for 4 h and then at room temperature overnight. The precipitate was collected by filtration, washed with acetone and dried to afford a light yellow solid (10.9 g, 21.5 mmol, 92%). The ^{1}H NMR data of compound **14** are identical to literature data[148]. MS (C18, ES$^+$) 505.85 $[M + H]^+$, $C_{21}H_{14}ClFIN_3O_0$ requires *M* 504.9854; ^{1}H NMR (500 MHz, DMSO) δ 11.70 (s, 1H), 9.34 (s, 1H), 8.95 (s, 1H), 8.36 (d, *J* = 8.8 Hz, 1H), 7.92 (d, *J* = 2.3 Hz, 1H), 7.76 (d, *J* = 8.7 Hz, 1H), 7.67 (dd, *J* = 8.9, 2.3 Hz, 1H), 7.47 (dd, *J* = 14.1, 7.9 Hz, 1H), 7.33 (dd, *J* = 14.5, 9.5 Hz, 3H), 7.19 (dd, *J* = 12.0, 5.2 Hz, 1H), 5.30 (s, 2H).

Compound 16[145]**:** In a flask containing compound **14** (10.5 g, 20.8 mmol) and compound **15** (3.75 g, 27 mmol) $Pd(PPh_3)_2Cl_2$ (695 mg, 5% mmol), DME (70 ml), EtOH (42 ml) and 2M Na_2CO_3 (70 mL) were added and the resulting mixture was heated at 70 °C for 2 h. The reaction was then cooled to room temperature and the product was precipitated as yellow solid. The solid was filetred, washed with water and dried under vacuum overnight to yield the pure product (7.5 g, 15.8 mmol, 76%). The ^{1}H NMR data for the compound **16** are identical to literature data[149]. MS (C18, ES$^+$) 474.01 $[M + H]^+$, $C_{26}H_{17}ClFN_3O_3$ requires *M* 473.0942; ^{1}H NMR (500 MHz, DMSO) δ 10.13 (s, 1H), 9.69 (s, 1H), 8.99 (s, 1H), 8.61 (s, 1H), 8.32 (dd, *J* = 8.7, 1.5 Hz, 1H), 8.00 (d, *J* = 2.5 Hz, 1H), 7.88 (d, *J* = 8.7 Hz, 1H), 7.76 (d, *J* = 3.7 Hz, 1H), 7.72 (dd, *J* = 8.9, 2.5 Hz, 1H), 7.48 (dd, *J* = 14.0, 8.0 Hz, 1H), 7.43 (d, *J* = 3.7 Hz, 1H), 7.37-7.25 (m, 3H), 7.25-7.12 (m, 1H), 5.27 (d, *J* = 7.9 Hz, 2H).

Compound 19[150]**:** To a solution of commercial available 2-mercaptoethanol **17** (7 g, 31.3 mmol) and compound **18** (2.19 ml, 31.3 mmol) in EtOH (70 ml) was added a solution of NaOMe in MeOH (30%, 7 ml) at room temperature. The mixture was stirred at room temperature overnight and then concentrated under reduced pressure. To the residue, water (60 ml) was added, followed by EtOAc (60 ml). The organic layer was washed with water and saturated brine, dried over anhydrous Na_2SO_4 and concentrated under reduced pressure. The crude product was purified by column chromatography (10-80% EtOAc in heptane) to give the title compound **19** as a colorless oil (5.75 g, 26

mmol, 83%). The ^{1}H NMR data for the compound **19** are identical to literature data[150]. MS (C18, ES$^+$) 222.10 [M + H]$^+$, $C_9H_{19}NO_3S$ requires *M* 221.1106; ^{1}H NMR (500 MHz, $CDCl_3$) δ 5.06 (br s, 1H), 3.73 (t, *J* = 6.0 Hz, 2H), 3.31 (d, *J* = 6.1 Hz, 2H), 2.72 (t, *J* = 6.0 Hz, 2H), 2.65 (t, *J* = 6.6 Hz, 2H), 1.43 (s, 9H).

Compound 20[150]**:** To a solution of compound **19** (5.7 g, 25.8 mmol) in DCM (250 mL) mCPBA (70%, 12.7 g, 51.6 mmol) was added at 0°C. The mixture was stirred at 0°C for 2 h and then aqueous $N_2S_2O_3$ solution was added to the reaction mixture, and the mixture was stirred at room temperature for 30 min. The mixture was extracted with EtOAc, and the organic layer was washed with saturated $NaHCO_3$, brine, dried over anhydrous Na_2SO_4 and concentrated under reduced pressure. The residue was purified by column chromatography (0-50% EtOAc in heptane) to give the title compound **20** as a colorless oil (4.31 g, 17.03 mmol, 66%). The ^{1}H NMR data for the compound (**20**) are identical to literature data[150]. MS (C18, ES$^+$) 254.40 [M + H]$^+$, $C_9H_{19}NO_5S$ requires *M* 253.0984; ^{1}H NMR (500 MHz, $CDCl_3$) δ 5.26 (br s, 1H), 4.12 – 4.10 (m, 2H), 3.64 (dd, *J* = 11.8, 5.9 Hz, 2H), 3.34 (t, *J* = 6.1 Hz, 2H), 3.25 (t, *J* = 6.1 Hz, 2H), 1.43 (s, 9H).

Compound 21[150]**:** 6N HCl (30 ml) was added to a solution of compound **20** (4.3 g, 16.98 mmol) in THF (105 ml) at room temperature. The mixture was then stirred at 60°C for 1.5 h, and concentrated under reduced pressure. EtOH was added to the residue and the mixture was concentrated again. The residue was purified by column chromatography (30-100% EtOAc in heptane) to give the title compound **21** as colorless oil (2.34 g, 15.3 mmol, 90%). The ^{1}H NMR data for the compound **21** are identical to literature data[150]. MS (C18, ES$^+$) 254.10 [M + H]$^+$, $C_4H_{11}NO_3S$ requires *M* 153.0460; ^{1}H NMR (500 MHz, DMSO) δ 5.37 (br s, 1H), 3.79 (t, *J* = 5.5 Hz, 2H), 3.54 – 3.51 (m, 2H), 3.38 – 3.36 (m, 2H), 3.22 – 3.19 (m, 2H).

Compound 22 (Modified Lapatinib)[19]**:** Compound **16** (6.33 g, 12.53 mmol) and compound **21** (2.3 g, 15.03 mmol) were dissolved in DCM (105 ml) containing molecular sieves (5Å) and the reaction mixture was heated at 40^{o}C for 3 h. Then, AcOH (43 ml) and NaBH $(OAc)_3$ (4.3 g, 16.98 mmol) were added and the mixture was stirred at room temperature overnight. After filtration to remove the molecular sieves, the

mixture was concentrated under reduced pressure and the residue was purified by column chromatography (0-20% MeOH in DCM) to give the title compound **22** as a yellow-orange solid (3.75 g, 6.13 mmol, 49%). R_f= 0.48 (10% MeOH in DCM). ^{1}H NMR (500 MHz, DMSO) δ 9.94 (s, 1H), 8.78 (s, 1H), 8.54 (s, 1H), 8.14 (dd, *J* = 8.7, 1.6 Hz, 1H), 8.04 (d, *J* = 2.5 Hz, 1H), 7.79 (m, 2H), 7.74 (dd, *J* = 9.0, 2.6 Hz, 1H), 7.46 (td, *J* = 8.0, 6.0 Hz, 1H), 7.35-7.10 (m, 4H), 7.06 (d, *J* = 3.3 Hz, 1H), 6.49 (d, *J* = 3.4 Hz, 1H), 5.25 (s, 2H), 4.02 (q, *J* = 7.1 Hz, 2H), 3.86 – 3.78 (m, 4H), 3.31 (dd, *J* = 13.0, 6.7 Hz, 2H), 3.02 (t, *J* = 6.7 Hz, 2H).; ^{13}C NMR (126 MHz, DMSO) δ 163.49, 161.79, 158.94, 153.07, 152.92, 152.40, 150.64, 149.08, 136.01, 131.06, 130.97, 129.15, 129.09, 124.72, 123.80, 123.78, 121.23, 116.79, 115.26, 115.04, 114.75, 114.60, 114.39, 110.01, 108.34, 69.86, 56.25, 55.39, 54.10, 45.62, 41.99; HRMS (ESI) 611.1521 $[M + H]^+$, $C_{30}H_{28}ClFN_4O_5S$ requires *M* 610.1526.

Compound 24 (triethylene glycol linker)[143]**:** To a solution of triethylene glycol di-*p*-tosylate **23** (4.52 g, 9.87 mmol) in anhydrous DMF (55 ml) di-tert-butyl iminodicarboxylate (**Boc$_2$NH**, 2.25 g, 10.86 mmol), Cs_2CO_3 (3.62 g, 11.8 mmol) and NaI (10% mol, 162 mg, 0.99 mmol) were added. The resulting mixture was stirred at 75 °C overnight and then left to cool to room temperature, whereby EtOAc (4 ml) and water (4 ml) were added. The mixture was transferred to a separating funnel and the organic layer was washed with 1N HCl, water and brine. The organic layer was dried over Na_2SO_4, filtered and concentrated *in vacuo* to give the crude product. Purification by column chromatography (20-50% EtOAc in heptane) gave compound **24** as colourless oil (1.32 g, 2.62 mmol, 27%). MS (C18, ES^+) 504.2 $[M + H]^+$, $C_{23}H_{37}NO_9S$ requires *M* 503.22; R_f= 0.52 (40% EtOAc in heptane). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.81 (d, *J* = 8.3 Hz, 1H), 7.36 (d, *J* = 8.0 Hz, 1H), 4.17 – 4.14 (m, 1H), 3.78 (t, *J* = 6.2 Hz, 1H), 3.70 – 3.66 (m, 1H), 3.58 (t, *J* = 6.1 Hz, 1H), 3.55 (s, 4H), 2.46 (s, 3H), 1.50 (s, 18H).; ^{13}C NMR (126 MHz, $CDCl_3$) δ 152.64, 144.82, 132.95, 129.83, 128.00, 82.33, 70.73, 70.13, 69.31, 69.23, 68.70, 45.12, 28.05, 21.65.

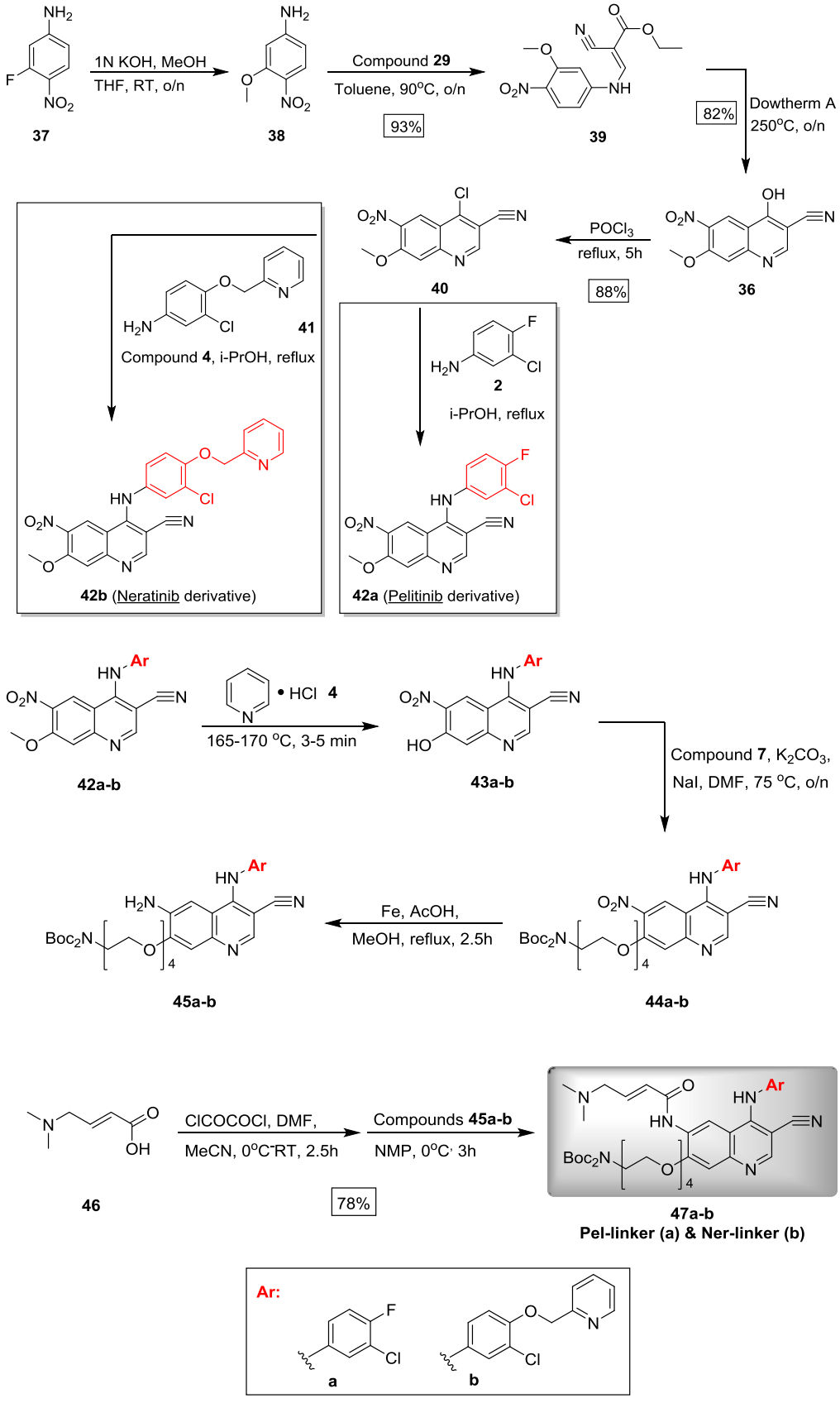
Compound 25 (Lap-linker): To a solution of compound **22** (1.32 g, 2.18 mmol) in anhydrous DMF (55 ml), NaH (60 % dispersion in mineral oil, 87.2 mg, 2.18 mmol) was added at 0^{o}C and stirred for 15 min. In another flask, compound **24** (1.3 g, 2.61

mmol) and NaI (10% mol, 36 mg, 0.99 mmol) were added. The resulting mixture of compound **24** was slowly added to compound **22** mixture at 0 °C and then left to warm to room temperature and stirred overnight. At this point, EtOAc (4 ml) and water (4 ml) were added. The organic layer was washed with 1N HCl, water and brine. The organic layer was dried over Na_2SO_4, filtered and concentrated *in vacuo* to give the crude product. Purification by column chromatography (0-30% MeOH in DCM) afforded compound **25** as dark yellow oil (860 mg, 0.91 mmol, 42%). R_f= 0.54 (10% MeOH in DCM).; ^{1}H NMR (500 MHz, DMSO) δ 9.86 (s, 1H), 8.73 (s, 1H), 8.56 (s, 1H), 8.16 (dd, *J* = 8.8, 1.5 Hz, 1H), 8.01 (d, *J* = 2.5 Hz, 1H), 7.80 (d, *J* = 8.8 Hz, 1H), 7.75 (dd, *J* = 9.0, 2.5 Hz, 1H), 7.50 – 7.42 (m, 1H), 7.31 – 7.28 (m, 3H), 7.20 (td, *J* = 8.6, 2.5 Hz, 1H), 7.07 (d, *J* = 3.2 Hz, 1H), 6.56 – 6.55 (d, *J* = 3.3 Hz, 1H), 5.27 (s, 2H), 5.13 (t, *J* = 4.8 Hz, 1H), 3.84 - 379 (m, 4H), 3.65 – 3.60 (m, 2H), 3.56 – 3.45 (m, 10H), 3.39 – 3.35 (m, 2H), 3.01 (t, *J* = 6.7 Hz, 2H), 2.71 (t, *J* = 5.7 Hz, 2H), 1.41 (s, 18H).; ^{13}C NMR (126 MHz, DMSO) δ 163.89 161.47, 158.05, 153.30, 152.58, 152.40, 150.27, 149.34, 140.16, 140.09, 133.50, 131.09, 131.00, 129.09, 128.97, 124.82, 123.82, 123.80, 123.02, 121.53, 117.00, 115.07, 114.78, 114.62, 114.40, 111.96, 108.27, 82.16, 70.13, 69.97, 69.86, 69.16, 68.83, 56.24, 55.46, 52.84, 51.77, 50.32, 47.24, 45.37, 28.01; HRMS (ESI) 942.3501 $[M + H]^+$, $C_{46}H_{57}ClFN_5O_{11}S$ requires *M* 941.3448;

3.1.2.3 Attempted syntheses of irreversible TKIs scaffold

Compound 27-36: Experimental procedures for compound **27**–**36** shown in ***Scheme 4-7*** and ***Scheme 4-8*** (both in ***Section 4.4.3***) are not included as the desired compounds were not synthesized and alternative compounds for the synthesis of the irreversible fluorescent inhibitors were synthesized. The procedures can be provided upon request.

3.1.2.4 Syntheses of Pel-linker and Ner-linker compounds



Scheme 3-3: Total synthesis of Pel-linker (47a) and Ner-linker (47b) compounds

Compound 38[151]**:** Compound **37** (10.5 g, 67.31 mmol) was dissolved in a 1:1 mixture of MeOH-THF (l80 ml), followed by addition of an excess of 1N KOH. The reaction mixture was heated under reflux overnight. The solution was cooled down to room temperature and treated with 10% aq HCl. After addition of EtAO the organic layer washed with water and brine, dried, filtered and concentrated. Purification by column chromatography afforded the compound **38**. The ^{1}H NMR data of compound **38** are identical to literature data[152]. MS (C18, ES$^+$) 168.9 $[M + H]^+$, $C_7H_8N_2O_3$ requires *M* 168.0535; ^{1}H NMR (500 MHz, DMSO) δ 7.82 – 7.79 (d, *J* = 9.1 Hz, 1H), 6.26 (s, 1H), 6.21 – 6.17 (dd, *J* = 9.0 Hz, 1.6 Hz, 1H), 3.81 (s, 3H).

Compound 39[138]**:** Compound **38** (9 g, 53.54 mmol) and compound **29** (18.1 g, 107.08 mmol) were dissolved in toluene (l00 ml) and refluxed overnight. The reaction mixture was then cooled down to 0^{o}C, filtered and washed with MTBE three times. The solid was dried to give compound **39** as a mixture of cis-trans isomers (14.5 g, 49.80 mmol, 93%) which was used without further purification. MS (C18, ES$^+$) 292.1 $[M + H]^+$, $C_{13}H_{13}N_3O_5$ requires *M* 291.0855.

Compound 36[138]**:** Compound **39** (13.5 g, 46.39 mmol) was added to refluxing (255^{o}C) Dowtherm A (375 ml) and the reaction mixture was refluxed for 4 h. After completion of the reaction, the mixture was cooled to room temperature and diluted with MTBE (350 ml), filtered and washed with MTBE three times. The solid was dried to give compound **36** (9.36 g, 38.20 mmol, 82%) which was used without further purification. The ^{1}H NMR data of compound **36** are identical to literature data[138]. MS (C18, ES$^+$) 245.3 $[M + H]^+$, $C_{11}H_7N_3O_4$ requires *M* 245.0437.; ^{1}H NMR (500 MHz, DMSO) δ 12.86 (s, 1H), 8.78 (s, 1H), 8.51 (s, 1H), 7.25 (s, 1H), 4.02 (s, 3H).

Compound 40[138]**:** Compound **36** (7.0 g, 28.57 mmol) was dissolved in $POCl_3$ (40 ml)) and stirred under refluxe until completion of the reaction (5 h). The reaction mixture was cooled to room temperature and slowly poured into ice/water (~400 ml) and left overnight. The solid was filtered off and washed with water to afford compound **40** (6.6 g, 25.10 mmol, 88%). The ^{1}H NMR data of compound **40** are identical to literature

data[138]. MS (C18, ES^+) 263.1 $[M + H]^+$, $C_{11}H_6ClN_3O_3$ requires *M* 263.0098.; ^1H NMR (500 MHz, DMSO) δ 9.26 (s, 1H), 8.83 (s, 1H), 7.89 (s, 1H), 4.13 (s, 3H).

Compound 42a[138]**:** In a flask containing compound **40** (2.8 g, 10.65 mmol) in *i*-PrOH (47 ml), 3-chloro-4-fluoro-aniline (**2**) (2.24 mg, 4.46 mmol) was added. The mixture was refluxed for 4 h and then cooled in an iced bath. The precipitate was collected by filtration, washed with chilled *i*-PrOH and dried to afford a white powder (3.15 g, 8.47 mmol, 80%) and the product was used to the next step without further purification. The ^1H NMR data for the compound **42a** are identical to literature data[153]. MS (C18, ES^+) 373.2 $[M + H]^+$, $C_{17}H_{10}ClFN_4O_3$ requires *M* 372.0425; ^1H NMR (500 MHz, DMSO) δ 9.34 (s, 1H), 8.90 (s, 1H), 7.74 – 7.70 (m, 2H), 7.58 – 7.41 (m, 2H), 4.10 (s, 3H).

Compound 42b[154]**:** A flask containing compound **40** (2.8 g, 10.65 mmol), aniline **41** (3.55 g, 15.13 mmol) and compound **4** (1.77 g, 15.13 mmol) in *i*-PrOH (45 ml) was refluxed for 3 h. This was then cooled in an iced bath and the precipitate was collected by filtration, washed with waster, MTBE and dried to afford the product which was used in the next step without further purification (4.63 g, 10.04 mmol, 94%). MS (C18, ES^+) 462.0 $[M + H]^+$, $C_{23}H_{16}ClN_5O_4$ requires *M* 461.0891; ^1H NMR (500 MHz, DMSO) δ 9.42 (s, 1H), 8.91 (s, 1H), 8.71 (d, *J* = 4.5 Hz, 1H), 8.10 (td, *J* = 7.8 Hz, 1.6 Hz, 1H), 7.76 – 7.72 (m, 2H), 7.62 (d, *J* = 2.0 Hz, 1H), 7.57 (dd, *J* = 6.8 Hz, 5.3 Hz, 1H), 7.43 – 7.34 (m, 2H), 5.43 (s, 2H), 4.10 (s, 3H).; ^{13}C NMR (126 MHz, DMSO) δ 159.71, 159.67, 159.39, 158.86, 158.00, 152.62, 144.86, 144.44, 136.67, 133.25, 131.74, 129.29, 127.87, 127.74, 126.88, 120.14, 119.58, 116.79, 112.61, 91.80, 75.23, 62.68.

Compound 43a: A flask containing compound **42a** (3.0 g, 8.07 mmol) and an excess of compound **4** was placed in a thermo-block pre-heated at 170°C. The progress of the reaction was monitored every min. Upon completion (5 min), the mixture was left to slowly cool to room temperature with slow addition of water. The precipitate was filtered, washed with water and dried to afford the crude product. Purification by column chromatography (0-40% EtOAc in heptane) afforded compound **43a** as a grey solid (2.65 mg, 7.40 mmol, 91%). R_f= 0.49 (15% EtOAc in heptane).; MS (C18, ES^+) 359.3 $[M + H]^+$, $C_{16}H_8ClFN_4O_3$ requires *M* 358.0269; ^1H NMR (500 MHz, DMSO) δ

9.28 (s, 1H), 8.83 (s, 1H), 7.70 (dd, *J* = 6.6 Hz, 2.4 Hz, 1H), 7.59 (s, 1H), 7.51 (d, *J* = 8.9 Hz, 1H), 7.46 – 7.39 (m, 1H).; ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.89, 154.45, 153.81, 143.68, 139.68, 128.34, 127.19, 127.08, 123.26, 117.95, 117.59, 115.69, 111.56, 106.35, 86.63.

Compound 43b: This compound was prepared from **42b** (4.5 g, 9.76 mmol) and pyridinium hydrochloride (**4**) using the method described above for **43a**. Purification by column chromatography (0-60% EtOAc in heptane) afforded compound **43b** as a grey solid (3.11 mg, 6.96 mmol, 71%). R_f = 0.45 (15% EtOAc in heptane).; MS (C18, ES^+) 448.2 $[M + H]^+$, $C_{22}H_{14}ClN_5O_4$ requires *M* 447.0734; ^{1}H NMR (500 MHz, DMSO) δ 9.26 (s, 1H), 8.77 (s, 1H), 8.64 – 8.62 (m, 1H), 7.83 (td, *J* = 7.7 Hz, 1.8 Hz, 1H), 7.63 – 7.58 (m, 2H), 7.53 (s, 1H), 7.42 (dd, *J* = 7.5 Hz, 5.0 Hz, 1H), 7.35 – 7.34 (m, 2H), 5.35 (s, 2H).; ^{13}C NMR (126 MHz, DMSO) δ 155.60, 155.52, 154.90, 153.65, 153.37, 148.72, 139.41, 138.91, 131.39, 128.74, 127.26, 124.10, 123.64, 122.59, 122.09, 114.83, 114.74, 110.86, 109.28, 86.16, 71.06.

Compound 44a: To a solution of compound **43a** (200 mg, 0.56 mmol) and K_2CO_3 (154 mg, 1.12 mmol) in anhydrous DMF (4 ml), tetraethylene glycol linker **7** (153 mg, 0.28 mmol) and a catalytic amount of NaI were added. The resulting mixture was stirred at 75°C overnight and then left to cool to room temperature, followed by addition of water (4 ml) and extraction with EtOAc (4 ml). The organic extracts were combined, dried over $MgSO_4$, filtered and concentrated *in vacuo* to give the crude product. Purification by column chromatography (00-50% EtOAc in heptane) gave the compound **44a** as yellowish oil (100 mg, 0.14 mmol, 25%). R_f = 0.59 (15% EtOAc in heptane).; MS (C18, ES^+) 734.4 $[M + H]^+$, $C_{16}H_8ClFN_4O_3$ requires *M* 733.2526; ^{1}H NMR (500 MHz, DMSO) δ 9.11 (s, 1H), 8.67 (s, 1H), 7.66 (s, 1H), 7.63 (dd, *J* = 6.9 Hz, 2.1 Hz, 1H), 7.53 – 7.46 (m, 1H), 7.40 – 7.34 (m, 1H), 4.48 – 4.44 (m, 2H), 3.85 – 3.82 (m, 2H), 3.66 – 3.60 (m, 4H), 3.55 – 3.46 (m, 8H), 1.43 (s, 18H).

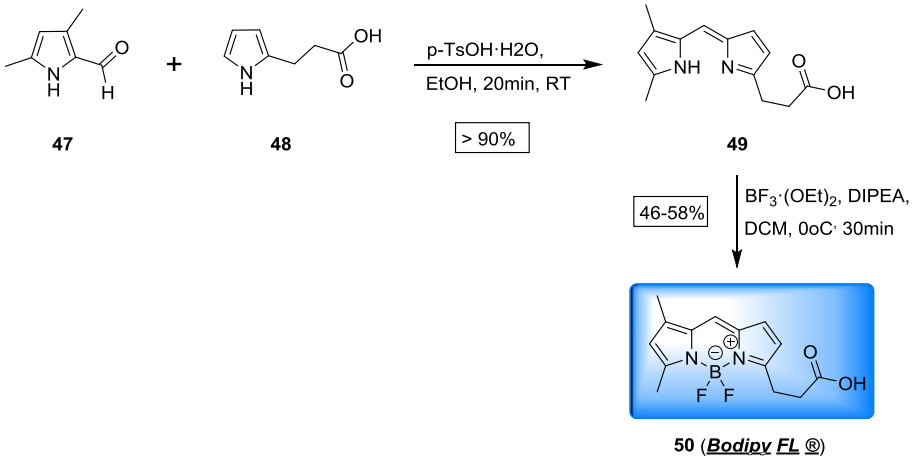
Compound 44b: This compound was prepared from **43b** (250 mg, 0.56 mmol), potassium carbonate (154 mg, 1.12 mmol), compound **7** (153 mg, 0.28 mmol) and a catalytic amount of NaI in anhydrous DMF (4 ml) using the method described above for

44a. Purification by column chromatography (00-50% EtOAc in heptane) gave the compound **44b** as yellowish oil (105 mg, 0.13 mmol, 23%). R_f = 0.48 (15% EtOAc in heptane).; MS (C18, ES^+) 823.5 $[M + H]^+$, $C_{40}H_{47}ClN_6O_{11}$ requires *M* 822.2991; ^1H NMR (500 MHz, DMSO) δ 9.13 (s, 1H), 8.63 – 8.60 (m, 2H), 7.92 – 7.85 (m, 1H), 7.65 (s, 1H), 7.59 (d, *J* = 7.5 Hz, 1H), 7.53 (s, 1H), 7.41 – 7.36 (m, 1H), 7.32 – 7.31 (m, 2H), 5.32 (s, 2H), 4.47 – 4.44 (t, *J* = 4.2 Hz, 2H), 3.85 – 3.82 (dd, *J* = 4.5 Hz, 3.8 Hz, 2H), 3.65 – 3.66 (m, 4H), 3.55 – 3.46 (m, 8H), 1.43 (s, 18H).

Compound 47a (Pel-linker): A flask containing compound **44a** (90 mg, 0.12 mmol), iron powder (21.5 mg), AcOH (43 µl, 0.76 mmol) and MeOH (1.2 ml) was refluxed for 2.5 h. After completion, the reaction mixture was cooled to room temperature and concentrated. The residue was dissolved in EtOAc-water (1:1) and filtrated. The organic layer was separated, washed with water, saturated $NaHCO_3$ and water again, dried over Na_2SO_4 and concentrated. The deprotected amine product (**45a**) was used to the next step without further purification. To a stirring solution of the acid **46** (57 mg, 0.44 mmol) and catalytic amount of DMF (1 drop) in MeCN (1.0 ml) at 0°C, oxalyl chloride ($(COCl)_2$ 37.3 µl, 0.44 mmol) was added. After stirring at 0°C for 30 min, the white suspension was left to stir at room temperature for a further 2 h. The mixture was then added dropwise to a solution of the deprotected amine **45a** in NMP at 0°C and the resulting mixture was stirred for 3 h at 0°C. The solvent was then removed and the residue was dissolved in EtOAc-sat. $NaHO_3$. The organic was dried over Na_2SO_4, filtered and concentrated *in vacuo* to give the crude product. Purification by column chromatography (20-100% EtOAc in heptane) gave compound **47a** as yellowish oil (76 mg, 0.094 mmol, 78%). R_f = 0.52 (10% MeOH in DCM).; ^1H NMR (500 MHz, DMSO) δ 9.10 (s, 1H), 8.55 (s, 1H), 8.47 (s, 1H), 7.27 (s, 1H), 7.15 (dd, *J* = 6.6 Hz, 2.2 Hz, 1H), 7.09 – 7.02 (m, 1H), 6.96 (m, 1H), 6.26 (d, *J* = 15.3 Hz, 1H), 4.34 – 4.31 (m, 2H), 3.97 – 3.93 (m, 2H), 3.77 – 3.66 (m, 6H), 3.59 – 3.54 (m, 6H), 3.12 (d, *J* = 5.6 Hz, 2H), 2.27 (s, 6H), 1.45 (s, 18H).; ^{13}C NMR (126 MHz, $CDCl_3$) δ 175.10, 164.16, 152.67, 152.36, 152.17, 151.49, 150.02, 147.26, 135.80, 134.99, 133.58, 126.59, 124.21, 121.37, 116.72, 113.72, 110.48, 109.16, 88.74, 82.35, 70.57, 70.20, 69.22, 69.02, 68.56, 60.18, 45.12, 28.04.; HRMS (ESI) 815.3527 $[M + H]^+$, $C_{40}H_{52}ClFN_6O_9$ requires *M* 814.3468.

Compound 47b (Ner-linker): This compound was prepared from compound **44b** (90 mg, 0.11 mmol), iron powder (21.5 mg), AcOH (43 µl, 0.76 mmol) and MeOH (1.2 ml) followed by addition of the acid **46** (57 mg, 0.44 mmol), catalytic amount of DMF (1 drop), oxalyl chloride (37.3 µl, 0.44 mmol) in MeCN (1.0 ml) and the deproetected amine **45b** in NMP using the method described above for **47a**. Purification by column chromatography (20-100% EtOAc in heptane) gave the compound **47b** as yellowish oil (70 mg, 0.077 mmol, 71%). R_f = 0.44 (10% MeOH in DCM).; ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.12 (s, 1H), 8.55 (dd, *J* = 4.1 Hz, 0.7 Hz, 1H), 8.46 (s, 1H), 7.76 – 7.69 (m, 1H), 7.63 – 7.61 (m, 1H), 7.31 – 7.19 (m, 3H), 7.08 – 6.91 (m, 3H), 6.27 (d, *J* = 15.3 Hz, 1H), 5.26 (s, 2H), 4.36 – 4.33 (m, 2H), 3.94 (dd, *J* = 5.1 Hz, 3.7 Hz, 2H), 3.76 – 3.64 (m, 6H), 3.61 – 3.53 (m, 6H), 3.17 (d, *J* = 6.0 Hz, 2H), 2.30 (s, 6H), 1.45 (s, 18H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 163.88, 156.53, 152.66, 151.82, 152.31, 150.45, 149.08, 147.14, 142.54, 137.02, 132.51, 128.59, 127.68, 126.19, 125.11, 123.56, 122.73, 121.25, 116.86, 113.84, 113.31, 109.37, 88.01, 82.32, 70.68, 70.61, 70.21, 69.23, 69.04, 68.65, 60.03, 45.28, 28.05.; HRMS (ESI) 904.3990 $[M + H]^+$, $C_{46}H_{58}ClN_7O_{10}$ requires *M* 903.3934.

3.1.2.5 Synthesis of Bodipy FL

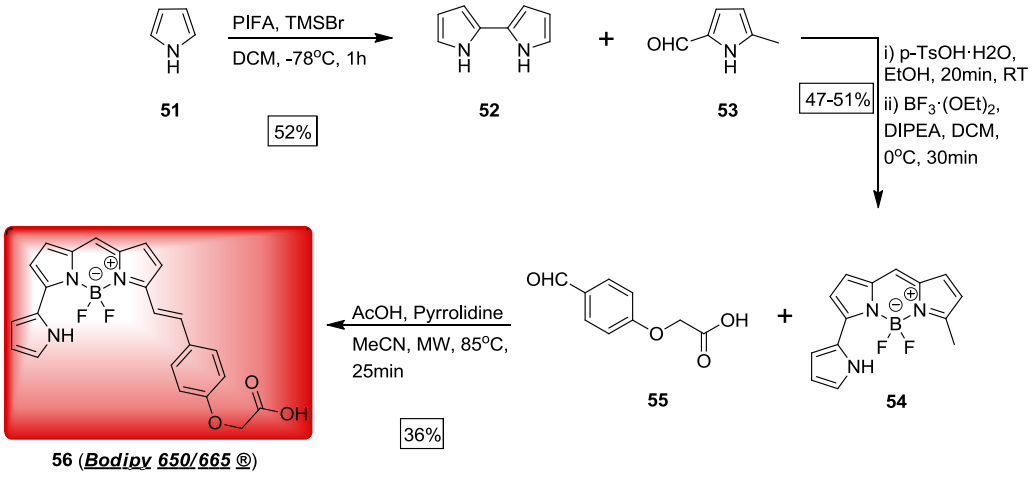


Scheme 3-4: Total synthesis of the "free" acid of Bodipy FL dye.

Compound 50 (Bodipy FL)[155]**:** Compound **47** (708 mg, 5.76 mmol) and the acid **48** (800mg, 5.76 mmol) were dissolved in ethanol (150 mL) and then pTsOH (1124 mg, 5.76 mmol) was added to the mixture. After 30 min at room temperature, the solvent

was removed *in vacuo*. The residue (compound **49**) was then dissolved in DCM (150 mL), followed by addition of DIPEA (5 mL, 28.8 mmol) and the resulting mixture was stirred for 5 min at room temperature. It was then cooled down to 0°C and boron trifluoride ethyl etherate was added (7 mL, 115.2 mmol). After stirring at this temperature for 1h, the mixture was left to slowly warm up to room temperature and stirred for another h. The solution was washed with water and brine, the organic layer was dried over MgSO4, filtered and concentrated *in vacuo* to give the crude product. Purification by column chromatography (0-20% MeOH in DCM) afforded the product as a red solid (771 mg, 2.65 mmol, 46%). The ^{1}H NMR data for the compound (**50**) are identical to literature data[156].; MS (C18, ES^{+}) 291.1 $[M]^{+}$, $C_{14}H_{15}BF_2N_2O_2$ requires M^{+} 291.1122; ^{1}H NMR (500 MHz, MeOD) δ 7.45 (s, 1H), 7.02 (d, *J* = 4.0 Hz, 1H), 6.33 (d, *J* = 3.9 Hz, 1H), 6.24 (s, 1H), 3.23 (t, *J* = 7.7 Hz, 2H), 2.77 (t, *J* = 7.7 Hz, 2H), 2.53 (s, 3H), 2.30 (s, 3H).

3.1.2.6 Synthesis of Bodipy 650/665



Scheme 3-5: Total synthesis of the "free" acid of Bodipy 650/665 dye.

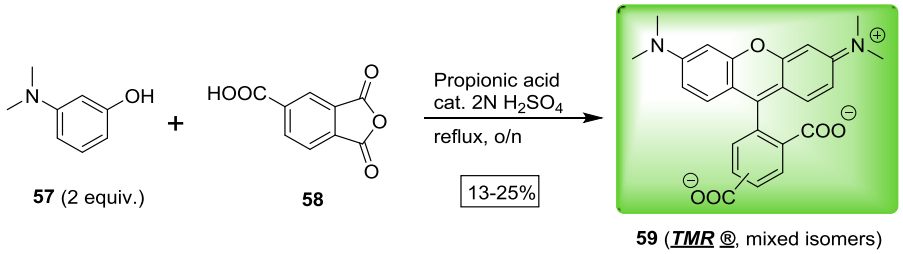
Compound 52 (bipyrrole)[157]**:** To a solution of pyrrole **51** (1.4 ml, 20.16 mmol) in DCM (300 ml) at -78°C, PIFA (2.9 gr, 6.74 mmol) and TMSBr (1.34 ml, 10.05 mmol) were added. The reaction mixture was stirred at -78 °C for 2 h before it was left to warm up to 0°C. It was then poured into ice-cold saturated $NaHCO_3$ (400 ml) and the resulting mixture was stirred at room temperature for 10 min. The organic phase was separated and the aqueous phase was extracted with DCM (2 x 350 ml). The combined

organic layers were dried over Na_2SO_4, filtered and concentrated *in vacuo* to give the crude product. Purification by column chromatography (0-50% EtOAc in heptane) afforded the product as a white powder (200 mg, 1.51 mmol, 52%) which was stored under N_2 at 4°C. The ^{1}H NMR data for the compound **52** are identical to literature data[157].; MS (C18, ES^+) 132.2 $[M]^+$, $C_8H_8N_2$ requires M^+ 132.1660; ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.21 (br s, 2H), 6.76 – 6.75 (m, 2H), 6.24 – 6.19 (m, 4H).

Compound 54: This compound was prepared from bipyrrole **52** (200 mg, 1.51 mmol) and compound **53** (157 mg, 1.51 mmol) using the method described above for compound **50**. Purification by column chromatography (0-40% EtOAc in heptane) afforded the product as a light brown solid (208 mg, 0.77 mmol, 51%).; R_f = 0.37 (20% EtOAc in heptane); MS (C18, ES^+) 271.2 $[M]^+$, $C_{14}H_{12}BF_2N_3$ requires M^+ 271.0778; ^{1}H NMR (500 MHz, MeOD) δ 7.18 – 7.12 (m, 4H), 6.96 – 6.91 (m, 2H), 6.35 (dd, *J* = 3.7 Hz, 2.7 Hz, 1H), 6.28 – 6.27 (d, *J* = 4.0 Hz, 1H), 2.59 (s, 3H).

Compound 56 (Bodipy 650/665)[158]: To microwave tube containing compound **54** (200 mg, 0.74 mmol, 1 equivalent) and the acid **55** (664 mg, 3.69 mmol, 5 equivalents in acetonitrile (3 mL) at 0°C, acetic acid (0.42 ml, 7.4 mmol, 10 equivalents and pyrrolidine (0.62 ml, 7.4 mmol, 10 equivalents were added. The tube was placed in a microwave oven and heated up at 85°C until completion of the reaction (25 min) using consutive 5-min steps. The reaction mixture was then concentrated *in vacuo* and purified by column chromatography (0-50% MeOH in DCM) to afford the product as a red solid (118 mg, 0.27 mmol, 36%).; R_f = 0.36 (10% MeOH in DCM); MS (C18, ES^+) 434.1 $[M + H]^+$, $C_{23}H_{18}BF_2N_3O_3$ requires M^+ 433.1409; ^{1}H NMR (500 MHz, MeOD) δ 7.45 (s, 1H), 7.03 – 7.01 (d, *J* = 4.0 Hz, 1H), 6.34 – 6.32 (d, *J* = 4.0 Hz, 1H), 6.24 (s, 1H), 3.26 – 3.20 (t, *J* = 7.6 Hz, 2H), 2.80 – 2.74 (t, *J* = 7.6 Hz, 2H), 2.53 (s, 3H), 2.30 (s, 3H).

3.1.2.7 Synthesis of tetramethyl 5/6-carboxy-rhodamine



Scheme 3-6: Total synthesis of the "free" acid of tetramethyl 5/6-carboxy-rhodamine dyes.

Compound 59 (tetramethyl 5/6-carboxy-rhodamine)[159]**:** N,N-Dimethyl-3-aminophenol **57** (3.0 g, 21.88 mmol) and 1,2,4-benzenetricarboxylic anhydride **58** (2.10 g, 10.95 mmol) were dissolved in a sealed tube containing propionic acid (10 mL), followed by addition of a catalytic amount of concentrated H_2SO_4 (2-3 drops). The reaction mixture was refluxed for 16 h and cooled to room temperature. The solvent was removed *in vacuo* to give the crude product as dark-purple slurry. Purification by column chromatography (10-40% MeOH in DCM) gave the mixture isomers of TMR (**59**) as a purple-gold solid (780 mg, 1.81 mmol, 17%). The ^{1}H NMR data for compound (**59**) are identical to literature data[160].; MS (C18, ES$^+$) 431.3 $[M]^+$, $C_{25}H_{23}N_2O_5$ requires M^+ 431.1601; ^{1}H NMR (500 MHz, MeOD) δ 8.79 (s, 0.5H), 8.26 (dt, J = 7.8, 1.8 Hz, 1H), 8.16 (d, J = 8.1 Hz, 0.5H), 7.86 (d, J = 1.4 Hz, 0.5H), 7.37 (d, J = 7.8 Hz, 0.5H), 7.27 (dd, J = 16.9, 9.5 Hz, 2H), 7.04 (dd, J = 9.5, 2.5 Hz, 2H), 6.93 (t, J = 2.0 Hz, 2H), 3.29 (s, 12H).

3.1.2.8 Coupling of TKIs with dyes

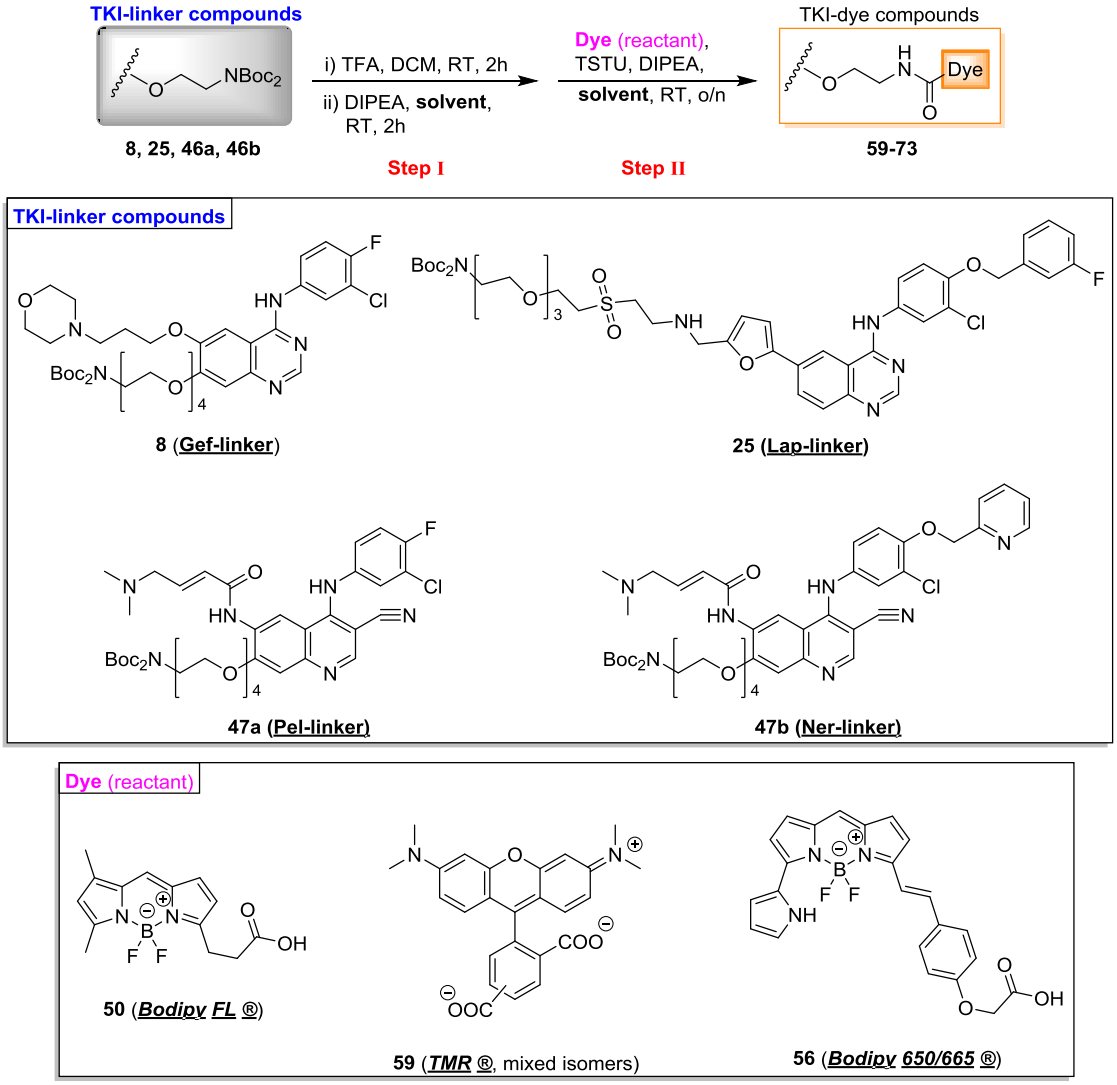


Table with conditions for step I and II

Dye (reactant) for step II	Solvent for step I and II	Product 59-73
Compound 50	DCM	TKI-Bod FL
Compound 59	DMSO or DMF	TKI-TMR5
		TKI-TMR6
Compound 56	DCM	TKI-Bod 650

Table with final fluorescent TKIs

Inhibitor	Fluorescent dye			
	Bodiby FL	TMR 5	TMR 6	Bodipy 650
Gefitinib	Compound 60	Compound 64	Compound 67	Compound 71
Lapatinib	Compound 61	Compound 65	Compound 68	Compound 72
Pelitinib	Compound 62	---	Compound 69	Compound 73
Neratinib	Compound 63	Compound 66	Compound 70	Compound 74

Scheme 3-7: Total synthesis of final fluorescent TKIs.

Dye-Inhibitor coupling_General Method (*Method A*): An excess of trifluoroacetic acid was (TFA) added to a solution of the tert-butyloxycarbonyl (BOC) protected TKI-linker compound (1 equivalent) in DCM and the reaction was stirred at room temperature for 1.5-2 h. The solvent was then removed using a steady stream of N_2 gas

or concentrated *in vacuo*, affording the deprotected compound (verified by LS-MS). It was then dissolved in solvent (DMF, DCM or DMSO), followed by addition of DIPEA (2 equivalents). In another flask (F2), the fluorescent dye (1 or more equivalents with respect to the TKI-linker compound) was dissolved in solvent (DMF, DCM or DMSO) and treated with TSTU (1.5 equivalents with respect to the dye) and DIPEA (1.5 equivalents with respect to the dye). The reaction mixture was stirred for 2 h at room temperature. The content of flask F2 was then added dropwise to the solution of the deprotected compound and the resulting mixture was stirred at room temperature overnight. The reaction mixture was then concentrated *in vacuo* and purified by preparative HPLC.

Compound 60 (Gef-Bod FL): This compound was prepared from compound **8** and compound **50** using *Method A* (described above). The crude material was purified by preparative HPLC (Boston C18 21*250mm 10µm, mobile phase: A: 0.1 % formic acid; B: acetonitrile) and the product was obtained as a green solid (31%). R_f = 0.50 (10% MeOH in DCM); ^{1}H NMR (500 MHz, MeOD) δ 8.45 (s, 1H), 8.04 (dd, *J* = 6.7 Hz, 2.6 Hz, 1H), 7.73 – 7.67 (m, 2H), 7.29 – 7.22 (m, 2H), 7.16 (s, 1H), 6.91 (d, *J* = 4.0 Hz, 1H), 6.24 – 6.23 (d, *J* = 4.0 Hz, 1H), 6.13 (s, 1H), 4.26 – 4.15 (m, 4H), 3.91 – 3.88 (m, 2H), 3.75 – 3.69 (m, 6H), 3.65 – 3.45 (m, 10H), 3.10 (t, *J* = 7.6 Hz, 2H), 2.76 – 2.70 (m, 2H), 2.52 (t, *J* = 7.7 Hz, 2H), 2.41 (s, 3H), 2.15 – 2.07 (m, 5H).; HRMS (ESI) 882.3720 $[M + H]^+$, $C_{43}H_{52}BClF_3N_7O_7$ requires *M* 881.3662.

Compound 61 (Lap-Bod FL): This compound was prepared from compound **25** and compound **50** using *Method A*. The crude material was purified by preparative HPLC (Boston C18 21*250mm 10µm, mobile phase: A: 0.1 % formic acid; B: acetonitrile) and the product was obtained as a green solid (31%). R_f = 0.45 (10% MeOH in DCM); ^{1}H NMR (500 MHz, MeOD) δ 8.68 (d, *J* = 1.7 Hz, 1H), 8.60 (s, 1H), 8.20 (dd, *J* = 8.8, 1.8 Hz, 1H), 7.98 (d, *J* = 2.6 Hz, 1H), 7.88 – 7.87 (d, *J* = 8.7 Hz, 1H), 7.70 (dd, *J* = 8.9, 2.6 Hz, 1H), 7.50 (td, *J* = 7.9, 5.9 Hz, 1H), 7.45 – 7.35 (m, 3H), 7.29 (s, 1H), 7.21 (d, *J* = 8.9 Hz, 1H), 7.15 (td, *J* = 8.5, 2.5 Hz, 1H), 7.00 – 6.99 (d, *J* = 3.4 Hz, 1H), 6.98 – 6.97 (d, *J* = 4.0 Hz, 1H), 6.58 – 6.57 (d, *J* = 3.4 Hz, 1H), 6.35 – 6.34 (d, *J* = 4.1 Hz, 1H), 6.19 (s, 1H), 5.32 (s, 2H), 3.97 (t, *J* = 5.7 Hz, 2H), 3.86 (s, 2H), 3.65 (t, *J* = 5.6 Hz,

2H), 3.61 – 3.58 (m, 6H), 3.51 (t, *J* = 5.3 Hz, 2H), 3.45 – 3.42 (m, 2H), 3.24 – 3.16 (m, 6H), 2.79 (t, *J* = 5.4 Hz, 2H), 2.56 (t, *J* = 7.7 Hz, 2H), 2.44 (s, 3H), 2.19 (s, 3H).; HRMS (ESI) 1016.3554 $[M + H]^+$, $C_{50}H_{54}BClF_3N_7O_8S$ requires *M* 1015.3488.

Compound 62 (Pel-Bod FL): This compound was prepared from compound **47a** and compound **50** using *Method A*. The crude material was purified by preparative HPLC (Boston C18 21*250mm 10µm, mobile phase: A: 0.1 % formic acid; B: acetonitrile) and the product was obtained as a green solid (27%). R_f = 0.21 (10% MeOH in DCM).; ^{1}H NMR (500 MHz, DMSO) δ 9.54 (s, 1H), 8.98 (s, 1H), 8.42 (s, 1H), 7.17 (s, 1H), 7.01 – 6.92 (m, 3H), 6.83 – 6.77 (m, 1H), 6.76 – 6.68 (m, 1H), 6.64 (d, *J* = 4.0 Hz, 1H), 6.05 (s, 1H), 5.98 (d, *J* = 3.8 Hz, 1H), 4.25 – 4.23 (m, 2H), 3.97 – 3.94 (m, 2H), 3.80 (d, *J* = 6.7 Hz, 2H), 3.75 – 3.59 (m, 6H), 3.50 – 3.47 (m, 2H), 3.36 (d, *J* = 5.1 Hz, 2H), 3.15 – 3.06 (m, 3H), 2.80 (s, 6H), 2.52 – 2.46 (m, 5H), 2.18 (s, 3H).; ^{13}C NMR (126 MHz, DMSO) δ 175.10, 164.16, 152.67, 152.36, 152.17, 151.49, 150.02, 147.26, 135.80, 134.99, 133.58, 126.59, 124.21, 121.37, 116.72, 113.72, 110.48, 109.16, 88.74, 82.35, 70.57, 70.20, 69.22, 69.02, 68.56, 60.18, 45.40, 45.12, 28.04.; HRMS (ESI) 815.3527 $[M + H]^+$, $C_{40}H_{52}ClFN_6O_9$ requires *M* 814.3468.

Compound 63 (Ner-Bod FL): This compound was prepared from compound **47b** and compound **50** using *Method A*. The crude material was purified by preparative HPLC (Boston C18 21*250mm 10µm, mobile phase: A: 0.1 % formic acid; B: acetonitrile) and the product was obtained as a green solid (31%). R_f = 0.25 (10% MeOH in DCM).; ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.04 (s, 1H), 8.52 (d, *J* = 4.3 Hz, 1H), 8.40 (s, 1H), 7.69 (td, *J* = 7.7 Hz, 1.7 Hz, 1H), 7.60 – 7.56 (m, 1H), 7.23 – 7.15 (m, 3H), 7.01 – 6.87 (m, 4H), 6.71 (d, *J* = 4.0 Hz, 1H), 6.25 (d, *J* = 15.4 Hz, 1H), 6.12 (d, *J* = 4.0 Hz, 1H), 6.01 (s, 1H), 5.22 (s, 2H), 4.26 (dd, *J* = 4.9 Hz, 3.4 Hz, 2H), 3.87 (dd, *J* = 5.1 Hz, 3.2 Hz, 2H), 3.67 – 3.50 (m, 8H), 3.46 – 3.42 (m, 2H), 3.46 – 3.42 (m, 2H), 3.36 – 3.31 (m, 2H), 3.19 – 3.12 (m, 4H), 2.53 – 2.46 (m, 5H), 2.30 (s, 6H), 2.14 (s, 3H).; HRMS (ESI) 1000.3877 $[M + Na]^+$, $C_{50}H_{55}BClF_2N_9O_7$ requires *M* 1000.3872.

Compound 64 & 67 (Gef-TMR5 & Gef-TMR6): These compounds were prepared from compound **8** and compound **59** (mixed 5- and 6-isomer of TMR) using *Method A*.

The products were separated and purified by preparative HPLC (Boston C18 21*250mm 10µm, mobile phase: A: 0.1 % formic acid; B: acetonitrile) to afford the pure **5-** (30%) and **6**-isomer (21%) as purple solids. **5-isomer (64)**: R_f = 0.28 (10% MeOH in DCM); ^{1}H NMR (500 MHz, MeOD) δ 8.62 (s, 1H), 8.43 (d, *J* = 10.8 Hz, 2H), 8.05 (d, *J* = 7.4 Hz, 2H), 7.75 (s, 1H), 7.37 (s, 1H), 7.29 (d, *J* = 7.8 Hz, 1H), 7.00 (dd, *J* = 25.7, 8.9 Hz, 4H), 6.61 (d, *J* = 9.0 Hz, 2H), 6.26 (s, 2H), 4.19 (s, 2H), 3.89 (s, 2H), 3.79 (d, *J* = 13.5 Hz, 2H), 3.78 – 3.72 (m, 8H), 3.71 (s, 4H), 3.66 (t, *J* = 11.7 Hz, 2H), 3.11 (s, 12H), 2.81 (d, *J* = 17.4 Hz, 6H), 1.98 (s, 2H).; HRMS (ESI): 1020.4073 $[M + H]^+$, $C_{54}H_{60}ClFN_7O_{10}$ requires *M* 1019.3996.; **6-isomer (67)**: R_f = 0.33 (10% MeOH in DCM); ^{1}H NMR (500 MHz, MeOD) δ 8.46 (s, 1H), 8.35 (s, 7H), 8.23 (d, *J* = 7.8 Hz, 1H), 8.10 (d, *J* = 7.2 Hz, 2H), 7.80 (s, 1H), 7.64 (s, 1H), 7.54 (s, 1H), 7.11 – 7.00 (m, 4H), 6.65 (d, *J* = 7.7 Hz, 2H), 6.35 (s, 2H), 4.25 (s, 2H), 3.89 (s, 2H), 3.81 (s, 4H), 3.71 (dd, *J* = 5.7, 3.3 Hz, 2H), 3.68 – 3.61 (m, 8H), 3.58 (t, *J* = 4.9 Hz, 2H), 3.15 (s, 12H), 2.92 (s, 4H), 2.89 – 2.82 (m, 2H), 2.04 (m, 2H).; HRMS (ESI) 1020.4067 $[M + H]^+$, $C_{54}H_{60}ClFN_7O_{10}$ requires *M* 1019.3996.

Compound 65 & 68 (Lap-TMR5 &Lap-TMR6): These compounds were prepared from compound **25** and compound **59** (mixed 5- and 6-isomer of TMR) using *Method A*. The products were separated and purified by preparative HPLC (Boston C18 21*250mm 10µm, mobile phase: A: 0.1 % formic acid; B: acetonitrile) to afford the pure **5-** (30%) and **6**-isomer (21%) as purple solids. **5-isomer (65)**: R_f = 0.36 (10% MeOH in DCM); ^{1}H NMR (500 MHz, MeOD) δ 9.87 (s, 1H), 8.85 (t, *J* = 5.5 Hz, 1H), 8.70 (s, 1H), 8.54 (s, 1H), 8.45 (s, 1H), 8.21 – 8.20 (dd, *J* = 8.1 Hz, 1.2 Hz, 1H), 8.13 – 8.11 (m, 1H), 8.02 – 8.01 (d, *J* = 2.6 Hz, 1H), 7.78 – 7.77 (d, *J* = 8.7 Hz, 1H), 7.75 – 7.73 (dd, *J* = 8.9, 2.5 Hz, 1H), 7.49 – 7.45 (m, 1H), 7.34 – 7.28 (m, 2H), 7.29 – 7.28 (d, *J* = 8.1 Hz, 1H), 7.26 (d, *J* = 9.0 Hz, 1H), 7.18 (td, *J* = 8.6, 2.6 Hz, 1H), 7.03 – 7.02 (d, *J* = 3.4 Hz, 1H), 6.53 – 6.52 (d, *J* = 3.4 Hz, 1H), 6.51 – 6.43 (m, 6H), 5.25 (s, 2H), 3.82 – 3.79 (m, 4H), 3.59 – 3.55 (m, 8H), 3.48 – 3.45 (m, 2H), 3.40 – 3.25 (m, 4H), 2.99 (t, *J* = 6.9 Hz, 2H), 2.92 (s, 12H), 2.70 (t, *J* = 5.6 Hz, 2H).; HRMS (ESI): 1154.3886 $[M + H]^+$, $C_{61}H_{61}ClFN_7O_{11}S$ requires *M* 1153.3822.; **6-isomer (68)**: R_f = 0.22 (10% MeOH in DCM); ^{1}H NMR (500 MHz, MeOD) δ 9.86 (s, 1H), 8.75 – 8.72 (m, 2H), 8.67 – 8.66 (m, 1H), 8.58 (s, 1H), 8.18 – 8.13 (m, 2H), 8.09 – 8.06 (m, 1H), 7.81 – 7.77 (m, 1H), 7.76 – 7.72 (dd, *J* = 9.3, 2.8 Hz, 1H), 7.65 (s, 1H), 7.50 – 7.44 (m, 1H), 7.37 – 7.27 (m,

3H), 7.23 – 7.16 (m, 1H), 7.04 (d, *J* = 3.7 Hz, 1H), 6.57 – 6.46 (m, 7H), 5.27 (s, 2H), 3.83 – 3.77 (m, 4H), 3.53 – 3.46 (m, 10H), 3.42 – 3.25 (m, 4H), 3.06 – 3.03 (m, 2H), 2.94 (s, 12H), 2.79 – 2.77 (m, 2H).; HRMS (ESI) 1154.3899 $[M + H]^+$, $C_{61}H_{61}ClFN_7O_{11}S$ requires *M* 1153.3822.

Compound 69 (Pel-TMR): This compound was prepared from compound **47a** and compound **59** using *Method A*. The products were separated and purified by preparative HPLC (Boston C18 21*250mm 10µm, mobile phase: A: 0.1 % formic acid; B: acetonitrile) to afford the pure product as purple solid (25%). R_f = 0.24 (40% MeOH in DCM).; ^{1}H NMR (500 MHz, DMSO) δ 8.33 (s, 1H), 8.57 (s, 1H), 8.31 (s, 1H), 8.03 (dd, *J* = 7.9 Hz, 1.7 Hz, 1H), 7.32 (d, *J* = 8.0 Hz, 1H), 7.26 – 7.19 (m, 1H), 7.14 – 8.03 (m, 5H), 6.77 – 6.70 (m, 4H), 6.59 – 6.49 (m, 2H), 5.98 (d, *J* = 3.8 Hz, 1H), 4.21 (m, 2H), 3.85 (m, 2H), 3.75 – 3.53 (m, 12H), 3.18 – 3.15 (m, 14H), 2.77 (s, 6H).; HRMS (ESI) 1027.3905 $[M + H]^+$, $C_{55}H_{56}ClFN_8O_9$ requires *M* 1026.3843.

Compound 66 & 70 (Ner-TMR): These compounds were prepared from compound **47b** and compound **59** (mixed 5- and 6-isomer of TMR) using *Method A*. The products were separated and purified by preparative HPLC (Boston C18 21*250mm 10µm, mobile phase: A: 0.1 % formic acid; B: acetonitrile) to afford the pure **5-** (27%) and **6-**isomer (21%) as purple solids. **5-isomer (66)**: R_f = 0.30 (30% MeOH in DCM).; ^{1}H NMR (500 MHz, MeOD) δ 8.63 (s, 1H), 8.46 – 8.44 (m, 2H), 8.10 (s, 1H), 7.96 – 7.92 (m, 1H), 7.82 – 7.76 (m, 1H), 7.60 – 7.56 (m, 1H), 7.30 – 7.08 (m, 4H), 7.03 – 6.93 (m, 5H), 6.70 – 6.66 (m, 2H), 6.57 (s, 2H), 6.43 (d, *J* = 15.6 Hz, 1H), 5.13 (s, 2H), 4.13 – 4.09 (m, 2H), 3.79 – 3.75 (m, 2H), 3.66 – 3.42 (m, 12H), 3.16 – 3.13 (m, 2H), 3.06 (s, 12H), 2.52 (s, 6H).; HRMS (ESI) 1116.4379 $[M + H]^+$, $C_{61}H_{62}ClFN_9O_{10}$ requires *M* 1115.4308. **6-isomer (70)**: R_f = 0.23 (30% MeOH in DCM).; ^{1}H NMR (500 MHz, MeOD) δ 8.70 (s, 1H), 8.46 (d, *J* = 5.0 Hz, 1H), 8.15 (s, 1H), 8.02 (d, *J* = 8.2 Hz, 1H), 7.95 (dd, *J* = 8.1 Hz, 1.6 Hz, 1H), 7.85 – 7.78 (m, 1H), 7.70 (s, 1H), 7.62 – 7.59 (m, 1H), 7.32 – 7.27 (m, 1H), 7.20 – 7.14 (m, 2H), 7.09 – 7.01 (m, 5H), 6.73 (dd, *J* = 9.5 Hz, 2.3 Hz, 2H), 6.63 – 6.62 (m, 2H), 6.51 – 6.45 (m, 1H), 5.15 (s, 2H), 4.18 – 4.15 (m, 2H), 3.81 – 3.78 (m, 2H), 3.60 – 3.42 (m, 12H), 3.17 – 3.13 (m, 2H), 3.08 (s, 12H), 2.55 (s, 6H).; HRMS (ESI) 1116.4379 $[M + H]^+$, $C_{61}H_{62}ClFN_9O_{10}$ requires *M* 1115.4308.

Compound 71 (Gef-Bod 650): This compound was prepared from compound **8** and compound **56** using *Method A*. The crude material was purified by preparative HPLC (Boston C18 21*250mm 10µm, mobile phase: A: 0.1 % formic acid; B: acetonitrile) and the product was obtained as a green solid (26%). R_f = 0.64 (20% MeOH in DCM); ^{1}H NMR (500 MHz, MeOD) δ 8.47 (s, 1H), 7.81 (dd, *J* = 6.6 Hz, 2.5 Hz, 1H), 7.55 – 7.52 (m, 1H), 7.45 (d, *J* = 8.7 Hz, 2H), 7.42 – 7.37 (m, 2H), 7.13 – 7.12 (m, 2H), 7.08 – 7.04 (m, 2H), 6.95 (d, *J* = 4.6 Hz, 1H), 6.93 (dd, *J* = 3.8 Hz, 1.4 Hz, 1H), 6.89 – 6.87 (m, 3H), 6.81 – 6.80 (d, *J* = 4.4 Hz, 2H), 6.74 – 6.73 (d, *J* = 4.4 Hz, 1H), 6.32 – 6.31 (m, 1H), 4.46 (s, 2H), 4.22 – 4.20 (m, 2H), 4.12 (t, *J* = 6.6 Hz, 2H), 3.90 – 3.88 (m, 2H), 3.72 (dd, *J* = 5.8 Hz, 3.7 Hz, 2H), 3.67 (t, *J* = 4.7 Hz, 2H), 3.65 – 3.63 (m, 2H), 3.60 – 3.53 (m, 6H), 3.49 – 3.46 (m, 2H), 2.52 (t, *J* = 7.2 Hz, 2H), 2.45 (s, 4H), 2.02 – 1.96 (m, 2H).; HRMS (ESI) 1023.3935 $[M + H]^+$, $C_{52}H_{55}BClF_3N_8O_8$ requires *M* 1022.3877.

Compound 72 (Lap-Bod 650): This compound was prepared from compound **25** and compound **56** using *Method A*. The crude material was purified by preparative HPLC (Boston C18 21*250mm 10µm, mobile phase: A: 0.1 % formic acid; B: acetonitrile) and the product was obtained as a green solid (26%). R_f = 0.64 (20% MeOH in DCM); ^{1}H NMR (500 MHz, MeOD) δ 8.47 (s, 1H), 8.41 (d, *J* = 1.7 Hz, 1H), 7.96 (dd, *J* = 8.7 Hz, 1.8 Hz, 1H), 7.76 (d, *J* = 2.4 Hz, 1H), 7.71 – 7.70 (d, *J* = 8.9 Hz, 1H), 7.50 (dd, *J* = 8.8 Hz, 2.5 Hz, 1H), 7.48 (s, 1H), 7.46 (d, *J* = 8.9 Hz, 1H), 7.43 (d, *J* = 17.1 Hz, 1H), 7.34 (td, *J* = 7.9 Hz, 5.9 Hz, 1H), 7.22 – 7.21 (d, *J* = 7.8 Hz, 1H), 7.18 (d, *J* = 9.6 Hz, 1H), 7.13 – 7.09 (m, 2H), 7.01 – 6.93 (m, 4H), 6.90 (s, 1H), 6.89 (d, *J* = 8.9 Hz, 2H), 6.83 – 6.81 (m, 2H), 6.75 – 6.74 (d, *J* = 4.0 Hz, 2H), 6.34 – 6.33 (d, *J* = 3.4 Hz, 1H), 6.31 – 6.30 (m, 1H), 5.11 (s, 2H), 4.46 (s, 2H), 3.94 (t, *J* = 5.5 Hz, 2H), 3.77 (s, 2H), 3.60 (t, *J* = 5.3 Hz, 2H), 3.57 – 3.54 (m, 6H), 3.48 – 3.46 (m, 2H), 3.38 (t, *J* = 6.9 Hz, 2H), 3.26 (t, *J* = 5.5 Hz, 2H), 3.18 – 3.16 (m, 2H), 2.72 (t, *J* = 5.3 Hz, 2H).; HRMS (ESI) 1157.3776 $[M + H]^+$, $C_{59}H_{57}BClF_3N_8O_9S$ requires *M* 1156.3703.

Compound 73 (Pel-Bod 650): This compound was prepared from compound **47a** and compound **56** using *Method A*. The crude material was purified by preparative HPLC (Boston C18 21*250mm 10µm, mobile phase: A: 0.1 % formic acid; B: acetonitrile)

and the product was obtained as a blue solid (30%). R_f = 0.36 (10% MeOH in DCM).; ^{1}H NMR (500 MHz, DMSO) δ 9.10 (s, 1H), 8.47 (s, 1H), 8.38 (s, 1H), 7.55 – 7.43 (m, 4H), 7.27 (s, 1H), 7.17 – 6.83 (m, 12H), 6.79 (d, *J* = 4.3 Hz, 1H), 6.35 – 6.32 (m, 1H), 6.22 (d, *J* = 15.3 Hz, 1H), 4.49 (s, 2H), 4.31 – 4.28 (m, 2H), 3.94 – 3.91 (m, 2H), 3.73 – 3.68 (m, 4H), 3.62 – 3.49 (m, 8H), 3.12 (d, *J* = 5.2 Hz, 2H), 2.77 (s, 6H).; HRMS (ESI) 1030.3783 $[M + H]^+$, $C_{40}H_{52}ClFN_6O_9$ requires *M* 1029.3723.

Compound 74 (Ner-Bod 650): This compound was prepared from compound **47b** and compound **56** using *Method A*. The crude material was purified by preparative HPLC (Boston C18 21*250mm 10µm, mobile phase: A: 0.1 % formic acid; B: acetonitrile) and the product was obtained as a blue solid (27%). R_f = 0.33 (10% MeOH in DCM).; ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.08 (s, 1H), 8.56 (d, *J* = 4.5 Hz, 1H), 8.45 (s, 1H), 7.76 – 7.70 (m, 1H), 7.63 – 7.60 (m, 1H), 7.55 – 7.44 (m, 3H), 7.27 (s, 1H), 7.21 – 7.07 (m, 5H), 7.01 – 6.82 (m, 9H), 6.78 – 6.77 (m, 1H), 6.25 (d, *J* = 15.5 Hz, 1H), 6.12 (m, 1H), 5.24 (s, 2H), 4.46 (s, 2H), 4.31 – 4.27 (m, 2H), 3.94 – 3.91 (m, 2H), 3.73 – 3.64 (m, 4H), 3.62 – 3.46 (m, 8H), 3.42 (d, *J* = 6.3 Hz, 2H), 2.50 (s, 6H).; HRMS (ESI) 1119.4249 $[M + H]^+$, $C_{59}H_{58}BClF_2N_{10}O_8$ requires *M* 1118.4189.

3.2 Determination of dissociation constants

Determination of the dissociation constant of compounds was performed in in DiscoveRx (San Diego, USA) using *KINOMEscan™ Profiling Service* technology. The protocol used is described below.

Kinase assay: For most assays, kinase-tagged T7 phage strains were prepared in an *E. coli* host derived from the BL21 strain. *E. coli* were grown to log-phase and infected with T7 phage and incubated with shaking at 32°C until lysis occured. The lysates were centrifuged and filtered to remove cell debris. The remaining kinases were produced in HEK-293 cells and subsequently tagged with DNA for qPCR detection. Streptavidin-coated magnetic beads were treated with biotinylated small molecule ligands for 30 min at room temperature to generate affinity resins for kinase assays. The liganded beads were blocked with excess biotin and washed with blocking buffer (SeaBlock (Pierce), 1% BSA, 0.05% Tween 20, 1 mM DTT) to remove unbound ligand and to reduce non-

specific binding. Binding reactions were assembled by combining kinases, liganded affinity beads, and test compounds in 1x binding buffer (20% SeaBlock, 0.17x PBS, 0.05% Tween 20, 6 mM DTT). All reactions were performed in polystyrene 96-well plates in a final volume of 0.135 ml. The assay plates were incubated at room temperature, shaking for 1 h and the affinity beads were washed with wash buffer (1x PBS, 0.05% Tween 20). The beads were then re-suspended in elution buffer (1xPBS, 0.05% Tween 20, 0.5 μM non-biotinylated affinity ligand) and incubated at room temperature, shaking for 30 min. The kinase concentration in the eluates was measured by qPCR.

3.3 Spectroscopic measurements of compound spectra

The excitation spectra were recorded using a Jasco V-630 UV-Vis spectrometer. The fluorescence excitation spectra of fluorescent inhibitors (between 100-500 nM in PBS) were recorded at various emission wavelengths depending on the fluorophore (Bodipy FL: 250 to 620 nm, TMR: 250 to 650 nm and Bodipy 650: 250 to 750 nm). The fluorescence of the solution was measured in a $1cm^3$ cuvette.

3.4 Glass passivation

Glass-bottom dishes used herein for seeding cells were treated with one of the following treatments as previously described[161].

Piranha cleaning: 30% w/v hydrogen peroxide was added to ice-cold concentrated sulphuric acid in 3:1 ration and left on ice for 5 min. The glass was treated with the mixture for 15 min at room temperature, rinsed with copious amount of water and left to dry overnight.

Basic cleaning: NaOH pellets were dissolved in deionised water to a final concentration of 7 M. The solution was further diluted to 0.1 M with deionised water. The glass was treated with 0.1 M NaOH for 1 h at room temperature, rinsed and left to dry overnight.

Bovine Serum Albumin (Sigma) was dissolved in PBS to a final concentration of 1% w/v and sterilized by filtration. The glass surface was coated with BSA for 1 h at room

temperature, rinsed and used promptly. BSA coating was applied after piranha or NaOH cleaning.

PEG-BSA nanogel: Nanogel was prepared as described by Zanetti-Domingues *et al.*[161] (and references therein)

For **single-molecule experiments with fluorescent inhibitors** (assessment of the non-specific binding of fTKIs), glass surfaces were treated with one of the aforementioned methods as indicated in section ***5.3.3***.

For **single-molecule experiments with ECD labelling**, glass surfaces were piranha cleaned and then coated with 1% BSA.

3.5 Cell Biology

All reagents, unless otherwise stated, were purchased from Invitrogen, UK. Cells were grown in 5% CO_2 in air at 37°C.

3.5.1 Cell culture

Chinese hamster ovary (CHO) cells stably transfected with wild-type EGFR or an EGFR mutant (L680N-EGFR, K721A-EGFR and Δ973-EGFR) were grown in phenol-red free Dulbecco's Modified Eagle medium (DMEM) supplemented with 10% (v/v) fetal bovine serum, 2mM glutamine, 1% penicillin-streptomycin, 100 μg/ml hygromycin and 100 μg/ml geneticin.

For **confocal experiments**, cells were seeded on uncoated 35 mm no. 1.5 glass-bottomed dishes (MatTek Corporation, USA) with 1.0 x 10^5 cells and with (stably transfected CHO cells) or without (transient transfected CHO cells) 250 ng/ml of doxycycline (~4 x10^5 receptors per cell) per sample. Cells were grown for two days before starvation.

For **single-molecule experiments**, cells were seeded on 35 mm no. 1.5 (high tolerance) glass-bottomed dishes (MatTek Corporation, USA) with 1.0 x 10^5 cells and with (stably transfected CHO cells) or without (transient transfected CHO cells) 50 ng/ml of

doxycycline (~5 x10^4 receptors per cell) per sample. Cells were grown for two days before starvation.

3.5.2 Transient transfection

Wild type CHO cells (ECCAC culture collection) were grown in phenol-red free DMEM/F12 nutrient mix supplemented with 10% (v/v) fetal bovine serum and 1% penicillin-streptomycin. For all experiments, a reverse transfection protocol was followed as described below. Briefly, 1 x 10^5 cells in 2 ml of media were transfected (before seeding in glass-bottomed dishes) with 1 μg of either EGFR mutants (L858R or L858R/T790M) or ErbB2-4 receptors DNA using ViaFect transfection reagent (Promega) at 4:1 reagent:DNA ratio.

3.5.3 Western blot

Either CHO cells stably expressing wild type EGFR (plus 250ng/ml of doxycycline hyclate (Sigma)) or wild type CHO cells reverse transfected with EGFR mutants were seeded on 6-well plates at a density of ~1.0 x 10^5 cells per well, grown to 80% confluency and serum-starved for 2 h (for stably expressing CHO cells, the same amount of doxycycline hyclate was added). Cells were treated with either the indicated concentrations of unlabelled murine-EGF (between 1 nM and 20 nM) or PBS 1x (negative control) for 2 h on ice. For the 9G8 nanobody[162] treatment, cells were treated for one h on ice at 4°C with 200 nM nanobody dissolved in PBS. After washing off the samples (3×ice-cold PBS), samples were lysed for 5 min on ice with 400 μl pre-heated lysis buffer (LDS buffer (Invitrogen) supplemented with 25 mM benzamidine, 1/100 Protease Inhibitor cocktail (Cell Signaling Technologies), 100 mM NaF, 1 mM Na_3VO_4 and 1% DTT). Lysates were cleared by centrifuging at 14000g for 5 min at 4°C.

Samples were run in parallel on 1.0 mm thick 3-8% Tris-Acetate NuPAGE gels (Invitrogen) with HiMark Prestained HMW and Novex Sharp Prestained protein standards (Invitrogen) using an XCell apparatus (Invitrogen). Proteins were blotted using an iBlot system (Invitrogen) on PVDF membranes, blocked for 1 h 4°C with 5% BSA in TBS+0.1% Tween and probed overnight with mouse anti-phosphotyrosine 4G10 antibody (Millipore). Gels were probed with secondary anti-mouse-HRP antibody

(Jackson ImmunoResearch) and incubated with Supersignal West Pico Chemiluminescent Substrate solution (Pierce) for 5 min, then imaged with a BioRad ChemiDoc MP system imager. Each blot was stripped with 25 ml stripping buffer (2% SDS, 0.75% β-mercaptoethanol, 62.5 mM Tris HCl pH 6.7) for 50 min at 60°C, and re-probed with an anti-EGFR cocktail composed of anti-EGFR D38B1 (Cell Signalling Technologies), anti-EGFR N-Terminal polyclonal ab137660 (abcam) and anti-EGFR polyclonal 10005: sc-03 (Santa Cruz Biotechnology), all derived from rabbit. Anti-rabbit-HRP (Jackson ImmunoResearch) was used for all blots and images were acquired as above.

Densitometry analysis was performed with ImageJ software (NIH)[163]. Bands were normalized to the amount of total EGFR.

3.5.4 Labelling of extracellular domain of receptors

EGF conjugates were custom-conjugated with NHS-ester or SE-ester dyes by Cambridge Research Biochemicals. Only 1:1 labelled fractions were used in Single-Molecule experiments. All **Affibodies and nanobodies** were conjugated in-house with maleimide dyes following manufacturers' instructions. **NRG1β** was conjugated in-house with maleimide dyes following manufacturers' instructions. **ErbB4 Adhiron** was conjugated in-house with NHS-ester dyes following manufacturers' instructions.

Table 3-1: List of ligands and dyes for the conjugates used in the present study

Ligand (*Manufacturer)*	**Dyes conjugated with ligand (*Manufacturer)***
EGF (*Peprotech*)	- Alexa 488 / *Molecular Probes Invitrogen* - CF640R / *Biotium*
Anti-EGFR Affibody (*Abcam*)	
Anti-EGFR nanobody EgB4 - ***gift by Dr Paul van Bergen en Henegouwen*** (***Utrecht University***)	Alexa 488 / *Molecular Probes Invitrogen*

Anti-ErbB2 Affibody ZErbB2:477 (m) (*Affibody Inc.*)	CF640R / *Biotium*
Anti-ErbB3 Affibody - ***gift by Dr J. Lofblom***	
NRG1β (*Peprotech*)	
Anti-ErbB4 Adhiron - ***gift by Dr D. Tomlinson*** (***University of Leeds***)	

3.5.5 Cell preparation and labelling

3.5.5.1 Permeability test of Fluorescent Tyrosine Kinase Inhibitors (fTKIs)

Wild type CHO cells were seeded on glass-bottomed dishes as described in section ***3.5.1***. After two days, the medium was changed to serum-free medium (phenol-red free DMEM/F12 nutrient mix supplemented with 1% penicillin-streptomycin) for 2 h. Cells were promptly imaged after replacing the medium with 200 nM of fTKIs dissolved in serum-free medium. For each fTKI, 20 confocal images of equatorial regions of the cells were collected over a period of 30 min from 2 replicates.

3.5.5.2 *In vitro* binding of fTKIs in live cells

CHO cells expressing wt EGFR were seeded on glass-bottomed dishes as described in section ***3.5.1***. After two days, the medium was washed thrice and replaced with a range of fTKIs' concentrations (50-400nM) dissolved in 0.1% serum medium (phenol-red free Dulbecco's Modified Eagle medium (DMEM) supplemented with 0.1% (v/v) fetal bovine serum, 2mM glutamine, 1% penicillin-streptomycin, 100 μg/ml hygromycin and 100 μg/ml geneticin) plus 250ng/ml of doxycycline hyclate for 2 h. After starvation, cells were washed thrice with PBS and cooled at 4°C on ice for 10 min. Cells were then labelled with 20 nM of either anti-EGFR Affibody or EGF (both labelled with either Alexa 488 or CF640R) for 1 h at 4°C to inhibit EGF-induced endocytosis. The N-

terminus of EGF was labelled at a 1:1 ratio by Cambridge Research Biochemicals (Cleveland, UK). The EGFR Affibody was labeled at a 1:1 ratio at its single cysteine residue. Cells were rinsed with ice-cold PBS and fixed with 3% paraformaldehyde plus 0.5% glutaraldehyde for 15 min at 4°C on ice, then for 15 min at room temperature. 10 confocal images per fTKI from 3 replicates were collected.

3.5.5.3 *In vitro* binding, binding specificity and dye effect of fTKIs and pre-fixed cells

CHO cells either stably expressing wt EGFR or transfected with any of ErbB2-4 were seeded on glass-bottomed dishes as described in section ***3.5.1***. After starvation, cells were washed thrice with PBS and cooled at 4°C for 10 min. For pre-treated cells with unlabelled Neratinib (Tocris Bioscience), cells were treated with 1μM of the commercial TKI during starvation and it was included during ECD labelling. Then, cells were labelled with 20 nM of an extracellular domain ligand (anti-EGFR Affibody, EGF, anti-ErbB2 or ErbB3 Affibody, NRG1β and anti-ErbB4 Adhiron, all labelled with either Alexa 488 or CF640R) for 1 h at 4°C. Cells were rinsed with ice-cold PBS and fixed with either 3% paraformaldehyde plus 0.5% glutaraldehyde or 3% paraformaldehyde for 15 min at 4°C, then 15 min at room temperature. Finally, cells were permeabilized with 0.05% Triton-X100 (Sigma) for 1 min. Cells were then labelled with fTKIs (100nM) dissolved in PBS just before imaging at room temperature and the labelling solution was remained in the dish during data acquisition.10 confocal images per fTKI from 3 replicates were collected using the microscope and set up described above.

3.5.5.4 Quantitative binding analysis

CHO cells stably expressing EGFR alleles (wt EGFR, L680N and Δ973) or wild type CHO transfected with L858R or L858R/T790M were seeded on glass-bottomed dishes as described in section ***3.5.1***. For Gefitinib pre-treatment, cells were treated with 1 μM unlabelled Gefitinib (Tocris Bioscience) during starvation and it was included during ECD labelling.

After starvation, cells were washed with PBS, cooled at 4°C for 10 min and then labelled for 1 h at 4°C with one of the following:

- 50 nM of anti-EGFR nanobody EgB4-Alexa 488 or 20 nM of anti-EGFR Affibody-CF640R (basal receptor)
- Either indicated concentration (1, 10 or 20 nM) of unlabelled EGF plus 50 nM of anti-EGFR nanobody EgB4-Alexa 488 plus or 20 nM of EGF-CF640R (EGF-stimulated receptor)
- 20 nM of anti-EGFR Affibody-CF640R plus 200 nM of 9G8 (9G8-treated receptor)

Cells were then rinsed with ice-cold PBS and fixed with 3% paraformaldehyde for 15 min at 4°C, then 15 min at room temperature. Finally, cells were permeabilized with 0.05% Triton-X100 (Sigma) for 1 min. Cells were then labelled with TMR labelled TKIs (100nM) dissolved in PBS just before imaging at room temperature and the labelling solution remained in the dish during data acquisition. Forty-five confocal images per fTKI from three replicates were collected.

3.5.5.5 Intradimer FRET between fTKIs

CHO cells expressing wt EGFR were seeded on glass-bottomed dishes as described in section ***3.5.1***. After starvation, cells were washed with PBS, cooled at 4°C for 30 min and then fixed with 3% paraformaldehyde for 15 min at 4°C, followed by 15 min at room temperature. Cells were then treated with Bodipy FL labelled TKIs (concentration between 10 and 400 nM, donor) dissolved in PBS and incubated for 20 min at room temperature. The donor labelling medium was then replaced with the acceptor medium (Bodipy 650 labelled TKIs at a concentration of either 0.2 or 2 μM). Data were collected without removing the acceptor labelling medium after incubation for 20 min. Twenty areas per fTKI from 2 replicates were collected.

3.5.5.6 Assessment of non-specific binding on glass surface

CHO cells expressing wt EGFR were seeded on glass-bottomed dishes as indicated in section ***5.3.3***. After starvation, cells were washed thrice with PBS, cooled at 4°C for 10 min and labelled for 1 h at 4°C with 4 nM of ECD labelled ligand (same as before) and 50 nM of fTKI (both dissolved in PBS). Cells were then rinsed with ice-cold PBS and fixed with 3% paraformaldehyde plus 0.5% glutaraldehyde for 15 min at 4°C, then 15 min at room temperature. Data were collected at room temperature.

3.5.5.7 Single-particle tracking with fTKIs

CHO cells expressing wt EGFR were seeded on glass-bottomed dishes as described in section ***3.5.1***. During starvation (2 h at 37°C) cells were treated with a mixture of labelled (20 nM) and unlabelled (980 nM) TKIs, both diluted in 0.1% medium. After starvation, cells were washed with 0.1% medium (without doxycycline), followed by ECD labelling using a mixture of unlabelled and labelled (either Alexa 488 or CF640R) ligand (Affibody or EGF). The final concentration of both was 100 nM with 10% of it being the labelled ligand. ECD labelling was performed at 37°C for 7 min.

3.5.5.8 Distance of Closest Approach experiments.

CHO cells stably expressing EGFR alleles (wt EGFR, L680N and Δ973) or wild type CHO transfected with L858R or L858R/T790M were seeded on glass-bottomed dishes as described in section ***3.5.1***. For Gefitinib pre-treatment, cells were treated with 1 μM unlabelled Gefitinib (Tocris Bioscience) during starvation and it was included during ECD labelling. After starvation, cells were labelled with 5 μM C18 Vybrant DiD or left in 0.1% medium (control), both without doxycycline, for 10-15 min at 37°C. After this step the same procedure described in section ***3.5.5.4*** was followed up to the fixation step. Data were collected at room temperature.

3.5.5.9 Characterisation of EgB4 and Affibody binding

CHO cells expressing wt EGFR were seeded on glass-bottomed dishes as described in section ***3.5.1***. After starvation, cells were washed with PBS, cooled at 4°C for 10 min and labelled for 1 h at 4°C with 50 nM anti-EGFR nanobody EgB4-Alexa 488 or 20 nM anti-EGFR Affibody-Alexa488 in the absence or the presence of 200 nM of 9G8 nanobody. Cells were rinsed and fixed with 3% paraformaldehyde plus 0.5% glutaraldehyde as described above. Thirteen confocal images of equatorial regions of the cells were collected from 3 replicates for concentration.

3.5.5.10 Single-particle tracking for colocalization experiments

CHO cells stably expressing EGFR alleles (wt EGFR, L680N and Δ973) or wild type CHO transfected with L858R or L858R/T790M were seeded on glass-bottomed dishes as described in section ***3.5.1***. For Gefitinib pre-treatment, cells were treated with 1μM unlabelled Gefitinib (Tocris Bioscience) during starvation and it was included during ECD labelling. After starvation, cells were washed thrice with 0.1% medium, followed by ECD labelling using a mixture of either affibodies with two different dyes (Alexa 488 and CF640R) or EGFs with the two same dyes. For 9G8 nanobody treatment, cells were treated for 30 min with 200 nM of 9G8 dissolved in 0.1% medium in the last 30 min of the starvation step before labelling. The final concentration of the Affibody mixture was 14 nM (1:1) and for EGF was 8 nM (1:1). For Affibody labelling, 10 min treatment at 37°C was used, whereas EGF mixture was added to cell culture dishes and cells were promptly imaged without rinsing steps, in order to minimise internalisation during labelling and imaging.

3.5.5.11 Fluorescence Localisation Imaging with Photobleaching (FLImP)

CHO cells stably expressing EGFR alleles (wt EGFR, L680N and Δ973) or wild type CHO transfected with L858R or L858R/T790M were seeded on glass-bottomed dishes as described in section ***3.5.1***. For Gefitinib pre-treatment, cells were treated with 1 μM unlabelled Gefitinib (Tocris Bioscience) during starvation and it was included during ECD labelling. After starvation, cells were washed thrice PBS, cooled at 4°C for 10 min and labelled for 1 h at 4°C with one of the following methods:

- 4 nM of anti-EGFR Affibody-CF640R (basal receptor)
- 4 nM of EG -CF640R (EGF-stimulated receptor)
- 4 nM of anti-EGFR Affibody-CF640R plus 200 nM of 9G8 (9G8-treated receptor)

Cells were then rinsed with ice-cold PBS and fixed with 3% paraformaldehyde plus 0.5% glutaraldehyde for 15 min at 4°C, then 15 min at room temperature. Data were collected at room temperature. 100-120 TIRF images per experiment from 3 replicates were collected.

3.6 Microscopy

3.6.1 Confocal Microscopy

Data were acquired at room temperature using a Leica TCS SP8 microscope with an oil immersion objective (NA=1.4, Leica) and Leica HyD hybrid detectors. Samples were illuminated with 488 nm, 550 nm or 640 nm light from an NKT Extreme supercontinuum light source and fluorescence was detected between 495 to 530 nm (for 488 nm excitation), 565 to 610 nm (for 550 nm excitation) and 650 to 690 nm (for 640 nm excitation). The pixel size of the confocal images was 76x76 nm.

3.6.1.1 Quantification of fTKIs binding

A machine learning algorithm (Trainable WEKA Segmentation[164]) plugin for ImageJ[165] was used to train a classifier that could reliably distinguish the membrane and non-membrane regions of the extracellular EGFR label. This classifier was applied to all of the imaged areas to exclude non-membrane pixels from the analysis. The pixel values of the ECD label channel (anti-EGFR Affibody, anti-EGFR nanobody or EGF) and the fTKI channel values were gathered and the fTKI intensity values were sorted by the corresponding Affibody channel value. The mean and standard deviation of the distribution of fTKI intensity values for each Affibody channel value was plotted.

3.6.1.2 Characterisation of EgB4 and Affibody binding

The fluorescence intensity values of pixels contributing to cell membranes were extracted to create a frequency distribution of membrane pixel intensities for each concentration of EgB4- and Affibody-Alexa 488. The median intensity value was plotted as a function of concentration, with error bars representing the upper and lower quartiles of the distribution. Frequency distributions of membrane pixel intensities between 9G8-treated and untreated labelled cells were compared.

3.6.1.3 Assessment of the expression of L858R and L858R/T790M EGFR mutants

The fluorescence intensity values of pixels contributing to cell membranes were extracted to create a frequency distribution of membrane pixel intensities for each

EGFR mutant labelled with Affy-CF640R. The median intensity value was plotted as a function of concentration, with error bars representing the upper and lower quartiles of the distribution. Frequency distributions of membrane pixel intensities between 9G8-treated and untreated labelled cells were compared.

3.6.2 FLIM-FRET Microscopy

Data were collected at room temperature using a Nikon Eclipse C1 confocal microscope with time-correlated single photon counting electronics (SPC830, Becker-Hickl GmbH). Samples were excited with a supercontinuum light source (Fianium SC450-4; 40 MHz repetition rate). Bodipy FL labelled TKIs (intradimer FRET between fTKIs) or ECD ligands with Alexa 488 (DOCA) were excited with 488 nm light and fluorescence was detected between 505 to 530 nm using a fast photomultiplier tube (PMC-100; Becker-Hickl GmbH). Confocal images of Bodipy 650 labelled TKIs (intradimer FRET between fTKIs) or Vybrant DiD (DOCA) were collected by excitation at 640 nm.

3.6.2.1 Lifetime measurements

Fluorescent intensity decays were best-fitted to a single exponential decay model where acceptor was absent and to a bi-exponential model when both donor and acceptor were present using SPCImage FLIM analysis software (Becker-Hickl GmbH). Donor lifetimes for FRET efficiency calculations were obtained by taking the mean of the distribution of the fluorescence lifetimes of pixels. The occurrence of FRET results in a decrease in the fluorescence lifetime of the donor τ_D in cells loaded with acceptor τ_DA. The FRET efficiency (E_{FRET}) was calculated from fluorescence lifetime data using the following formula

$$E_{FRET} = 1 - \frac{\tau_D}{\tau_{DA}}$$

Intradimer FRET between fTKIs. For each cell in the FLIM images, regions of interest were drawn to isolate the membrane. Lifetime distribution and the mean lifetime and standard deviation from these pixels were used to for the plots in ***Section 5.3.2*** and also to calculate a mean FRET efficiency.

Distance of Closest Approach experiments. For each cell in the FLIM images, regions of interest were drawn to isolate the membrane and the mean lifetime from these pixels was used to calculate a mean FRET efficiency. The same region of interest was then applied to the corresponding image of Vybrant DiD labelling to determine the corresponding mean acceptor intensity. The mean acceptor intensity was converted to a density with units of acceptors per R_0^2 using a calibration factor previously determined from samples with a known DOCA[136]. A model obtained from Monte Carlo simulations of a single donor at different distances above a plane of acceptors was fitted to each dataset to estimate the ensemble averaged DOCA for that experiment.

3.6.3 Single-Molecule Microscopy

Data for all methods used were acquired using Axiovert 200M microscopes with TIRF slider (Zeiss, UK) with a NA=1.45 oil immersion objective (100x, α-Plan-Fluar, Zeiss, UK) and an EMCCD camera (Andor, UK)[166]. The microscopes are also equipped with a wrap-around incubator (Pecon XL S1).

For **single-particle tracking (SPT)** (two-colour experiments), samples were excited with a Lighthub-6 combiner (Omicron) which contains solid state lasers emitting at 488 nm (60 mW, Phoxx, Omicron), 561 nm (75 mW, Jive, Cobolt) and 642 nm (140 mW, Phoxx, Omicron), which are co-aligned into a polarisation-maintaining single mode optical fibre. Samples were imaged at 37^{o}C acquiring an image every 0.05 sec for a period of 30 sec. Two-color TIRF images of the basolateral surfaces of cells were chromatically separated by a beam splitter[166].

For **fluorophore localisation imaging with photobleaching (FLImP)** data, samples were illuminated with 640 nm light taken from a diode laser (100mW, Cube, Coherent) from a fibre-coupled laser combiner (Andor) or VortranStradus 638 nm diode laser (Laser Technology, Inc., USA). Images were collected at room temperature and at 3.57 Hz for a period of ~140-200 sec, until most of the single-molecule spots had been photobleached.

3.6.3.1 Data analysis

Identification of spots, extraction of single particle tracks and registration of two-colour TIRF images (for SPT experiments) in order to map the relative positions of the probes over the time course of data acquisition were made using custom Bayesian algorithms reported by Rolfe *et al*[96].

Calculation of diffusion coefficient. Mean squared displacement (MSD) curves were calculated as $MSD(\Delta T)=\langle|r_i(T+\Delta T)-r_i(T)|^2\rangle$ where $|r_i(T+\Delta T)-r_i(T)|$ is the displacement between position of track i at time T and time $T+\Delta T$ and the average value is over all pairs of points separated by ΔT in each track. The average instantaneous diffusion coefficient (D) for these tracks was calculated by fitting a straight line to the first 2 points of the MSD curve then calculating D directly from the gradient m of the fit, $D=m/4$. The errors in the MSD curve were calculated by repeating the MSD curve calculation 200 times, each time on a different synthetic dataset created by randomly resampling with replacement the time points within each track, the tracks present within each dataset, and the datasets present (bootstrap resampling). The error in D, $\sigma(D)$, is calculated from the standard deviation of the D fits obtained from each bootstrap-resampled MSD curve.

Colocalisation analysis. The fraction of co-localized tracks was reported as tracks in which a receptor particle in one channel colocalizes with another particle in the other channel, spending at least three 50 ms frames in total moving together within a pixel of each other[137]. Coincidental colocalization statistics were calculated for a dataset as follows. In each channel, a randomized set of tracks of the same size as the measured set was produced, where each track in the random sample was chosen (with replacement) from the measured tracks for that channel, recentred at random with uniform probability density across the field of view, rotated through a random angle with uniform probability density between 0 and 360 degrees, and randomly flipped in x with probability 0.5. The random tracks therefore have key properties such as durations and path structures representative of the true tracks, but now randomized in distribution and orientation. The colocalisation statistics were then calculated for the randomized tracks. This was performed a total of 50 times for each dataset, and the colocalisation statistics pooled to give a final estimate coincidental colocalisation fraction for that dataset. The reported colocalisation fraction is then the fraction for the real data minus the estimated coincident colocalisation fraction. The **duration of individual events** in which a track

in one channel moves within a pixel of a track in the other channel and then they move apart again was also calculated. To reduce the impact of localization error on these results a temporal Gaussian smoothing filter of FWHM 4 frames (200ms) was applied to the position traces before the colocalisation analyses.

FLImP distribution. FLImP analysis is described in detail in Needham *et al.* (2013)[70] and in Zanetti-Domingues *et al.* (2015)[71] but will also be briefly explained here.

The FLImP analysis workflow is shown in ***Figure 3-1***. From total internal reflection fluorescence (TIRF) images a two-fluorophore spot is selected. A potentially suitable, for FLImP analysis, spot have the following characteristics: two intensity levels of equal height separated by a photobleaching step, the lowest level decays to zero intensity after photobleaching and each level contains more than 10 frames of data[70,71]. In such a spot when one fluorophore bleaches, the centroid position of the spot shifts. A global least-squares 7-parameter fit is used to identify the best intensity, x and y positions for each fluorophore (6 parameters) and the full width at half-maximum of the point spread function(1 parameter), from which the separation between the molecules ($r = \sqrt{(x_1 - x_2)^2 = (y_1 - y_2)^2}$) can be calculated with a precision determined by the localization error. Monte Carlo bootstrap method is used to assess the probability distribution of the model parameters and the x and y fluorophore positions converted to a distribution of separations. Empirical posterior FLImP distributions were obtained based on discrete fluorophore separation measurements that had confidence intervals of less than 6-7 nm.

Calculation of the proportion of FLImP measurements in the tethered-extended dimer range (see ***Figure 6-15***). The full FLImP distribution is shown in grey. The FLImP distributions for subsets of the measurements whose 69% confidence intervals overlap with the ranges of Affibody separations of the expected tethered (4.1 [-1.9 +4.2] nm) or extended (12.0 ± 1.6 nm) are shown in green and yellow respectively. Affibody separations >12.0+1.6 nm are highlighted in blue. The proportion of the measurements corresponding to each of the subsets was calculated, with error bars calculated by bootstrap-resampling the data 1000 times with replacement and repeating the analysis.

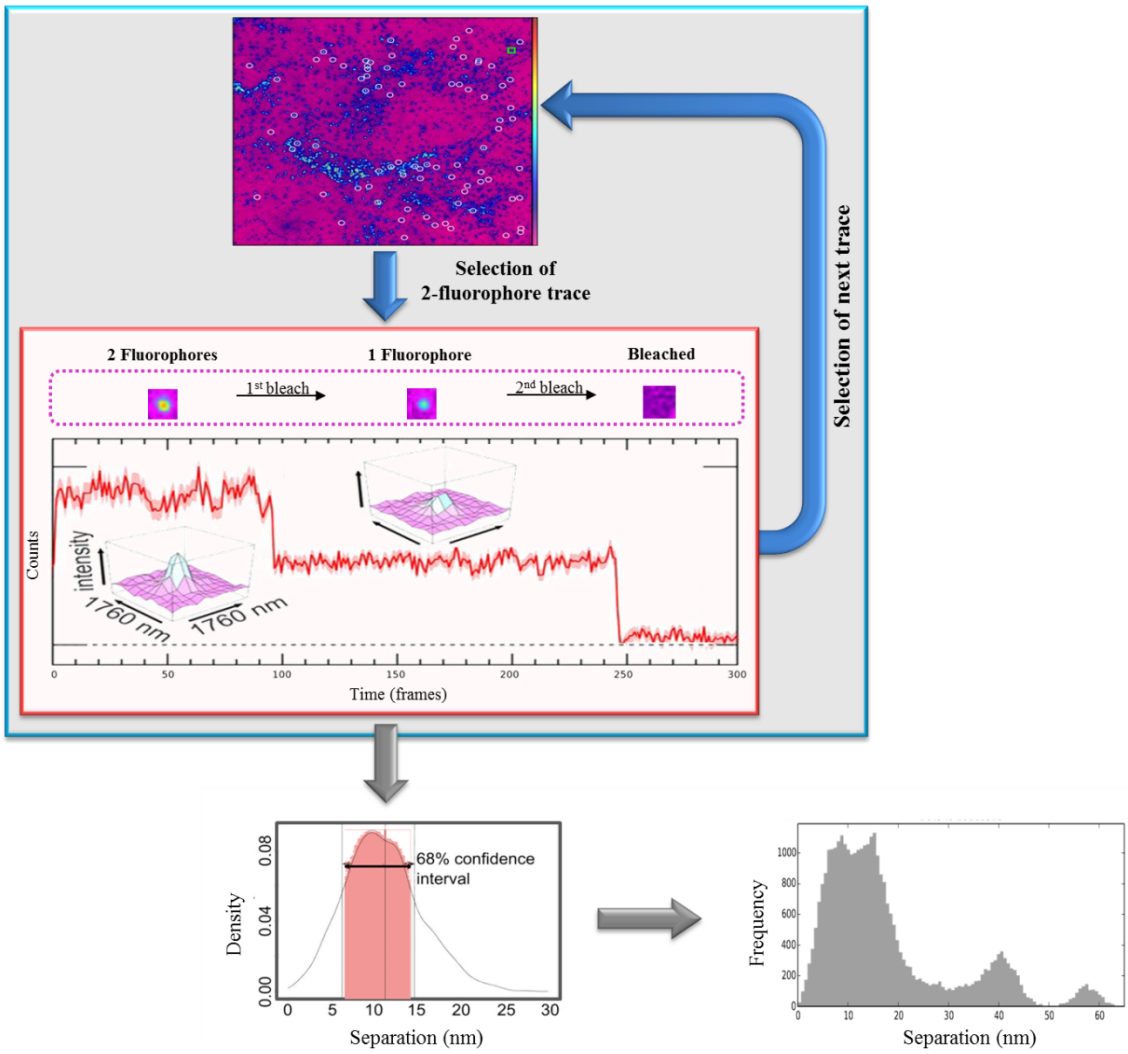


Figure 3-1: Schematic representation of FLImP analysis workflow.

4 Design and Synthesis of Fluorescent Tyrosine Kinase Inhibitors

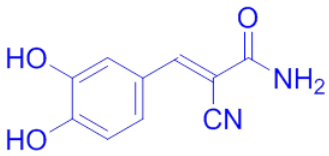
4.1 Introduction

In order to develop small-molecule probes for successfully labelling the ICDs of EGFR family, the reactive warhead of the probes discussed in ***Section 1.4.1*** must be able to recognise the target in a selective manner. Over the years, numerous small-molecule TKIs have been developed to target and inhibit the TKDs of EGFR family receptors with modest to great selectivity and affinity. Therefore, TKIs could be used as the recognition element which would afford probes with dual capabilities: probing the structure of the receptors and studying TKIs' properties in a cellular context.

The most crucial step towards the development of fluorescent tyrosine kinase inhibitors (fTKIs) is to choose TKIs that are very selective and exhibit high affinity to their target(s). To do this, understanding their binding mode and reviewing the available literature is indispensable. Moreover, insight on TKIs would also allow identification of a position on their structures where the linker can be attached without affecting the binding. Lastly, the selection of cell-permeable and photostable fluorescent dyes for tagging the TKIs is also important and will be discussed.

4.2 TKIs for EGFR Family

The paramount role of the TKD in EGFR activation was discovered in 1987 from parallel mutational studies on the ATP binding site which showed the absolute dependency on the TKD for tyrosine phosphorylation activity[167,168]. A year later, the first compound (***Figure 4-1***), inspired by natural products[169], that had the ability to inhibit EGFR phosphorylation by binding to TKD was published[170]. This discovery gave rise not only to the development of countless compounds for EGFR but also to drug discovery for many other protein kinases.



Tyrphostin AG 99

Figure 4-1: Structure of the first published TKI for EGFR.

The need for increase selectivity and potency of EGFR TKIs led to a much more complex structure of today's inhibitors compared to Tyrphostin. TKIs for the EGFR family are ATP-competitive inhibitors which target the ATP binding site located between the N- and C- lobe in TKD. The quinazoline scaffold mimics the adenine base of ATP and exploits similar interactions within the cleft (***Figure 4-2***). Diversity around the quinazoline core confers selectivity to the target and increases potency for efficient displacement of ATP, characteristics which are essential to reduce the side effects of treatment.

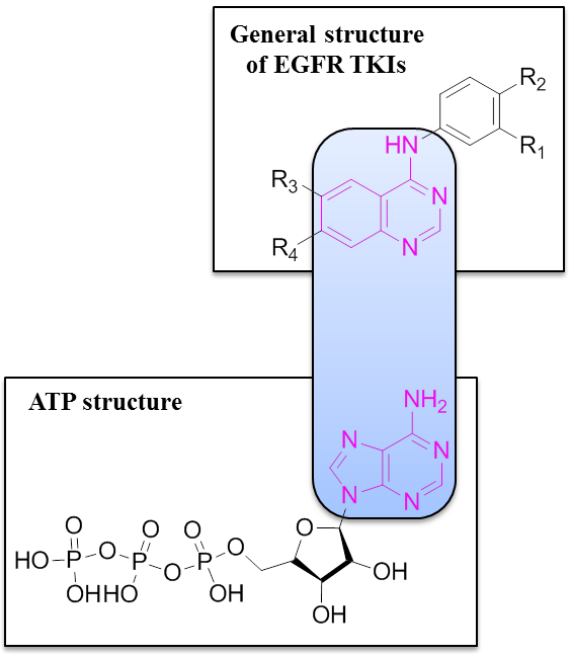


Figure 4-2: Common characteristics between ATP and EGFR TKIs. The quinazoline core (highlighted area) of EGFR TKIs mimics the adenine base of ATP and exploits the same interactions within the ATP-cleft.

Kinase inhibitors[22], are classified by their generation (first, second and third) and each generation differs from each other in evolved structural characteristics for

improving effectiveness[22,171] (***Figure 4-3***). The first generation of EGFR family inhibitors includes reversible TKIs such as Gefitinib, Lapatinib and Erlotinib whereas the second generation comprises irreversible (covalent) inhibitors and it emerged as a solution for overcoming the resistance of EGFR mutants (T790M) to the treatment with inhibitors of the first generation[172]. Third generation TKIs were designed to selectively inhibit T790M mutation over wild-type (wt) as the irreversible ones, apart from Afatinib[173], were unsuccessful in clinical trials. EGFR TKIs can be further classified by their type (I and II). Type I inhibitors such as Gefitinib (first generation), Pelitinib and Afatinib (both second generation), target[22] and promote[65] the active TKD conformation whereas Type II (or Type I½[174]) such as Lapatinib and Neratinib (first and second generation respectively) target[22] and promote[65] the inactive (further discussion about Type I and II in ***Section 5.3.1***). The only difference between the two types lies in the *para*-substituent on the aniline ring (see ***Section 4.3***).

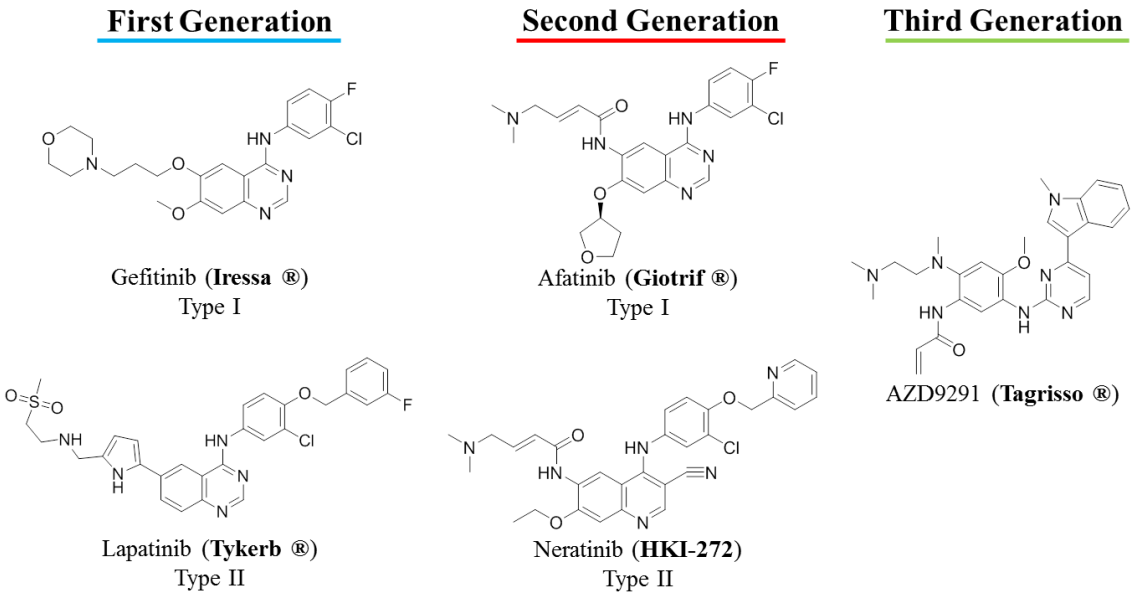


Figure 4-3: Chemical structures of some TKIs and their generation. EGFR TKIs can be divided in three generation: reversible (first), irreversible (second) and mutant-selective inhibitors (third). Inhibitors can be further classified in type I (target inactive KD) and type II (target active KD) compounds.

Apart from a few exceptions such as TAK-285[57] and AEE778[175], the structure of the core of both first and second generation TKIs is the same (anilinoquinazoline scaffold). The chemical structure of the third generation TKIs differs from the rest (only the pyrimidine core is retained), and from each other, as they have been developed in order

to circumvent resistance of the, so-called, "gatekeeper" mutation (T790M) which precludes the rigid anilinoquinazoline core from slotting in the ATP-binding cleft.

Even though ErbB family members are highly homologous, selective binding of some TKIs against either EGFR or ErbB2 has been reported. The FDA approved Gefitinib and Erlotinib are EGFR-selective compounds[176], whereas as ARRY-380[177] and CP-724,714[178] are ErbB2-selective. The rest drugs, either approved or in clinical trial, are multi-targeted TKIs with examples of dual (EGFR/ErbB2)[47] inhibitors, equipotent against 3 members (EGFR/ErbB2 and 4[179] or EGFR/ErbB2-3[180]) or even compounds that bind to all four ErbBs[181].

Based on the reported potencies and selectivity profiles of the published compounds as well as from studies investigating the binding affinities of various compounds[176,182,183] I have created a map showing the structure-selectivity relationship of first and second generation TKIs (***Figure 4-4***).

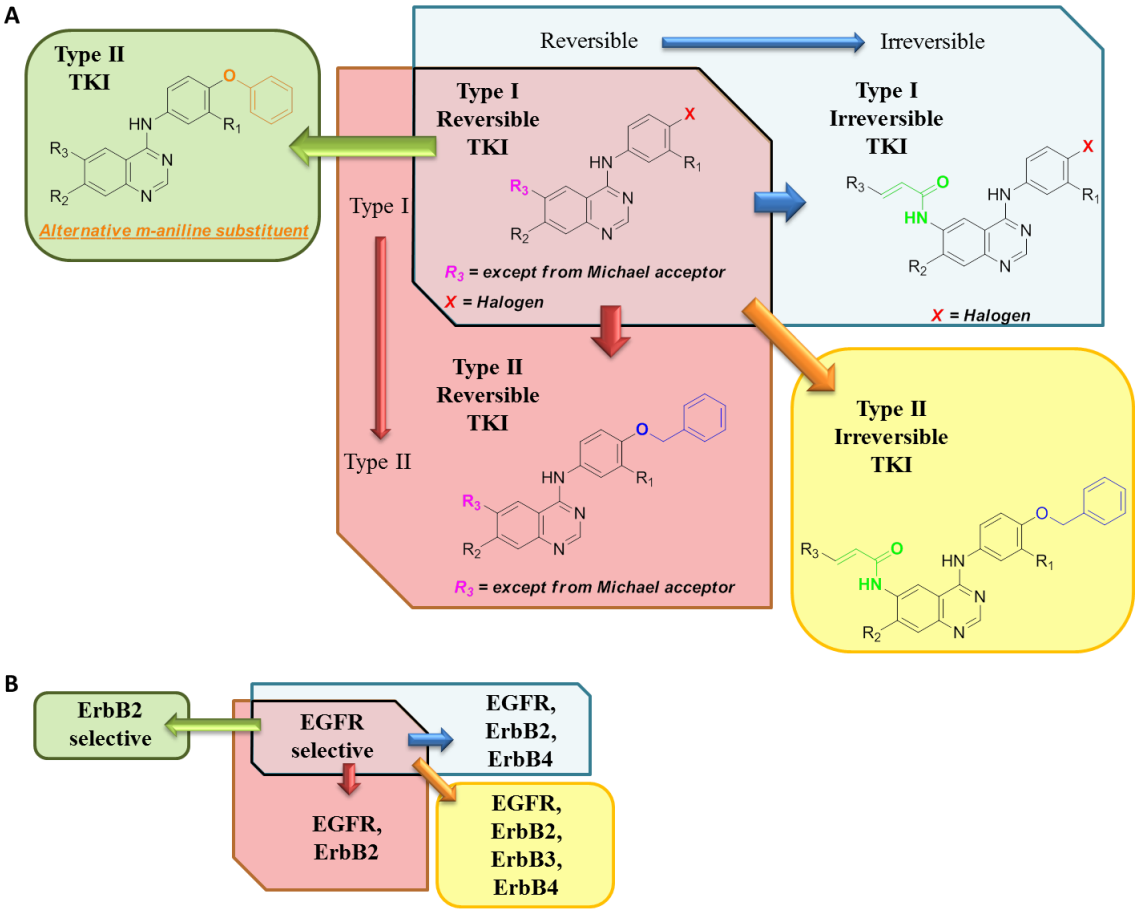


Figure 4-4: Structure-selectivity relationship map of TKIs against ErbB family members.

Type I reversible (first generation) TKIs are EGFR-selective with affinities against the other members which are at least 200-fold higher. Moving to type II by substitution of the small halogen with a benzyl (or other 5- or 6-member ring) alcohol (red arrow) results in dual EGFR/ErbB2 inhibitors. The affinity against ErbB4 is also increased but not at the same levels as for EGFR and ErbB2. A similar increase in potency is also found when comparing Type I reversible and irreversible TKIs (blue arrow), this time however, the affinity against all three receptors (EGFR, ErbB2 and 4) is at low nanomolar levels. Replacement of both structural features (which leads to type II irreversible inhibitors – orange arrow) produces compounds which can bind all ErbB members with similar potency. The last observation is related to the two ErbB2-selective compounds mentioned earlier. While the increase in ErbB2 affinity is expected when the halogen is replaced by a bulkier group (green arrow), the decrease affinity against EGFR is in contrast to what was observed when moving from reversible type I to type II. The answer may lie in the more rigid phenol substituent of those two inhibitors whereas the extra carbon on the benzyl alcohol ring present in type II reversible inhibitors allows it to freely rotate around itself.

4.3 Identification of Labelling Position and Selection of TKIs for fTKIs development

After having understood the structural features that define selectivity but also having reviewed the most potent compounds (characteristics essential for the recognition element), I then moved to the linker part of the probes' structure. In order to successfully design the fluorescent versions of TKIs it is essential to incorporate the linker (and subsequently the fluorescent dye) in a position with no or minimal effect on the binding affinity of the inhibitor.

The binding mode of first and second generation inhibitors, which share the anilinoquinazoline core, is common and can be divided into three sections (***Figure 4-5A***): the quinazoline core (red) and the aniline substituent (purple) which are placed with the ATP-binding site and the R_3 and R_4 substituents (blue) which point out of the pocket. The potency of the inhibitors is mainly due to hydrogen bond interactions between the nitrogen atoms (N) of the **quinazoline core** and nearby residues within the hydrophobic pocket[184]. Each ErbBs inhibitor can form up to three different hydrogen

bonds, depending on its atoms and its orientation. Furthermore, crystal structures showed that one of these interactions, between N_3 and a threonine residue within the binding pocket (Thr766), is bridged by a water molecule as the distance between them is longer than the minimum required to form a direct hydrogen bond. To overcome the requirement for the presence of water for this interaction, scientists from Wyeth developed two inhibitors, Pelitinib (EKB-569)[138] and Neratinib (or HKI-272)[181], in which the N_3 atom was ingeniously replaced by a carbon atom bearing a cyano-group in order to minimize the distance between the nitrogen and Thr766[185]. Apart from the three aforementioned interactions, Peng *et al.*[186] recently reported an additional hydrogen bond interaction (do not occur in anilinoquinazoline-based inhibitors). This novel interaction takes place between the Asp residue of the DGF motif, outside the pocket, and an amine on some inhibitors' scaffold which contributed to a 50-fold higher potency compared to inhibitors lacking this interaction.

The **aniline ring**, which plays a role in the selectivity of TKIs towards ErbB2-4 as discussed earlier, is buried deep into the cavity and interact with other residues that stabilize the binding. These interactions however are weaker (hydrophobic, π-π and Van der Waals interactions) and hence their contribution to the total potency is less important compared to hydrogen bond interactions. While the contribution of the solvent-exposed **R_4 substituent** to the binding is also minimal, **R_3 substituent** is a key element of the irreversible TKIs.

Second-generation irreversible inhibitors have been designed to form a covalent adduct with EGFR, ErB2 and ErbB4 (ErbB3 lacks the required residue in the specified position), by exploiting a cysteine residue (Cys773 for EGFR, Cys805 for ErbB2 and Cys803 for ErbB4) which resides just outside of the pocket (***Figure 4-5B***). For this, the substituent on the R_3 position of the reversible inhibitors was replaced with a Michael acceptor, which is usually an α,β-unsaturated amide.

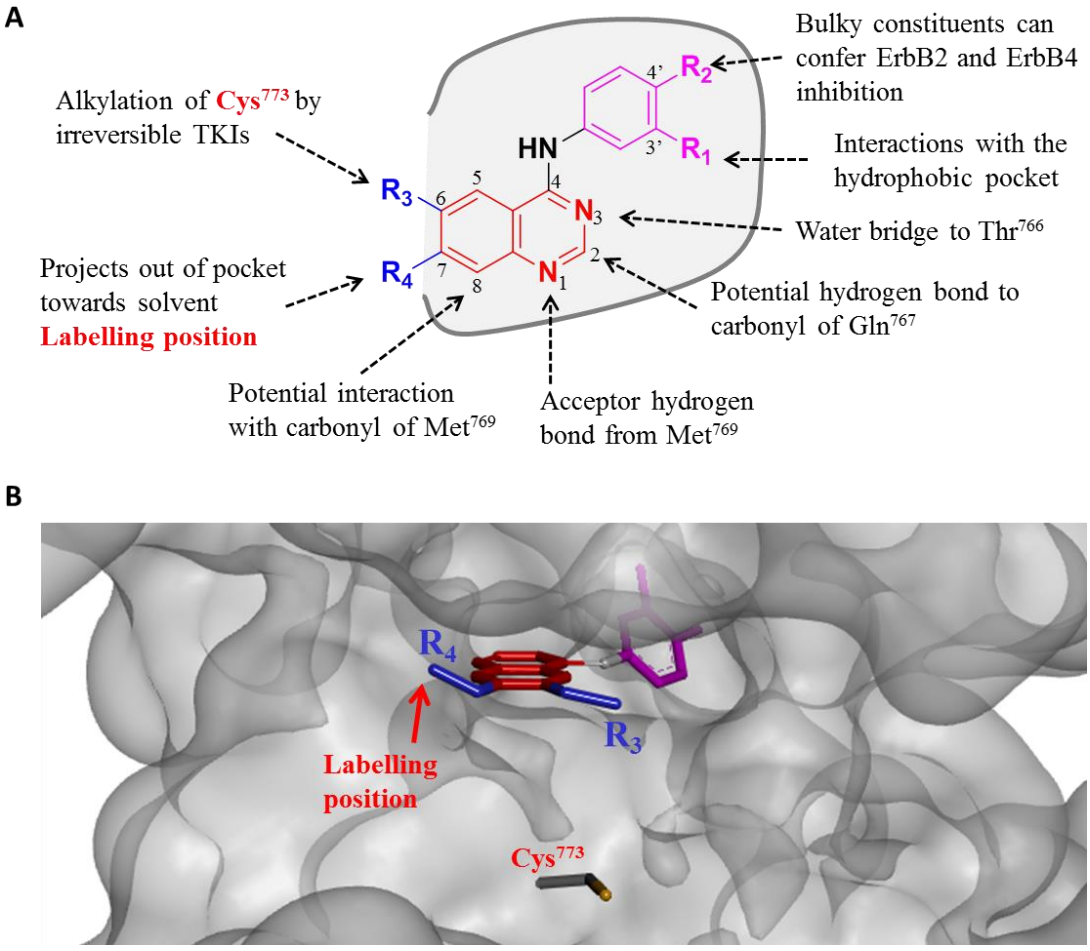


Figure 4-5: Diagram of common interactions of TKIs within the ATP cleft of the kinase domain (A) and 3-dimensional view of a TKI in the binding site (B).

The above analysis, together with the examination of all the available crystal structures of different anilinoquinazoline-based inhibitors bound to ErbB members clearly suggested that the most suitable position for attaching the linker would be on the R_4 group. R4 substituents points out of the pocket, do not participate in essential interactions and in many TKIs the functional group at this place is a methoxy or an ethoxy group. This also guided me in the selection of the water-soluble polyethylene glycol (PEG) as a linker which would act as an extension of the existing ether groups, if present in the parent TKIs. Additionally, PEG linkers are commercially available in different lengths and have been used in similar cases[124,143].

Of equal importance as the position for the incorporation of linker, and for the same reasons, is the length of the linker between the dye and the inhibitor. The ideal linker length was calculated by performing molecular modelling (Dr Martyn Winn, STFC) using the PDB structure of 2ITY (***Figure 4-6***). After having screened three different

linkers, the tetra-ethylene glycol linker was found to give the optimum distance. This occurs when the linker is long enough in order for the dye not to alter the binding of the inhibitor but not so long as to result in a very flexible dye. Flexible dyes can potentially be a problem when low-nanometer precision is required (single-molecule localization experiments) or when protein-protein interactions are monitored with FRET (within ~10 nm distance[83]).

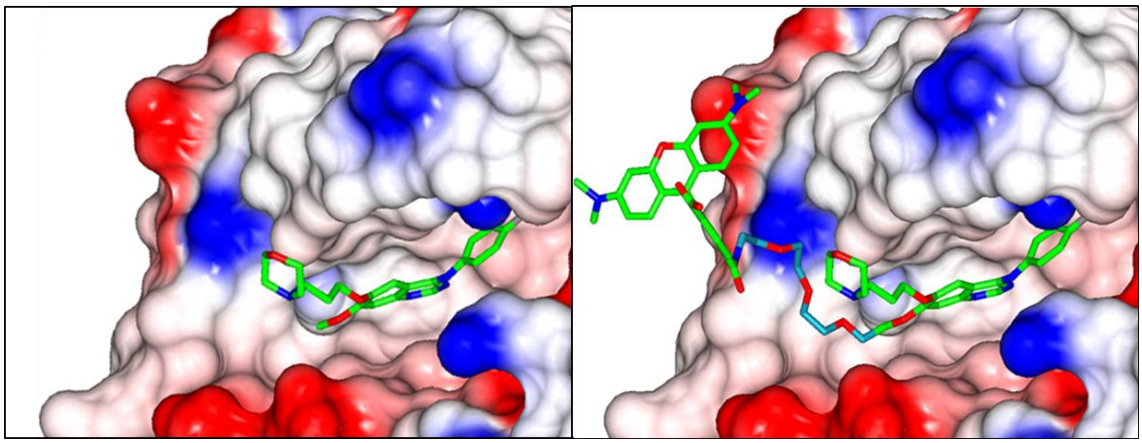


Figure 4-6: Molecular modelling for determination of optimum linker length.

The last component of the fTKI structure is the fluorescent dye. Due to their low molecular weight (MW), small-molecule probe properties such as permeability, wash-out kinetics and intracellular distribution have been shown to be greatly dependent on the choice of the dye[187,188]. The importance of dyes on dye-EGFR ligand conjugates, as well as their fluorescent properties and their suitability for single-molecule studies, have been also explored in depth by my co-workers[104,161]. Even though these studies provide useful information regarding dyes that one should avoid, depending on the experiment, the properties of a novel low-MW conjugate cannot be accurately pre-determined based on their results. Cases where a certain dye did not affect the permeability of a probe but instead affected other conjugates with different warheads[189–191] confirm the difficulty in predicting the physicochemical properties of a molecule. In view of this fact it was decided to select (based on the hydrophobicity and cell permeability of dyes-small-molecule conjugates) three different fluorescent dyes (green, orange and far-red) for labelling the compounds in the last step of the synthesis (***Figure 4-7***). The variety of different dye colours would allow me not only to choose the best conjugate for a certain experiment but also to perform FRET experiments and 2-/3-colour imaging.

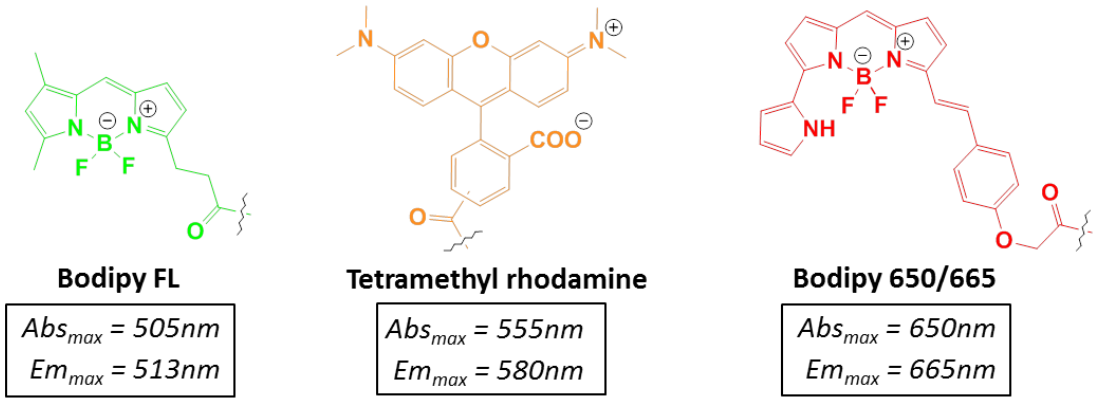


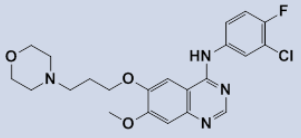
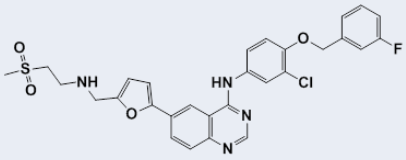
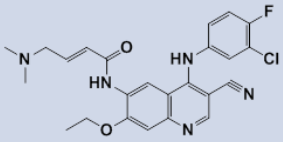
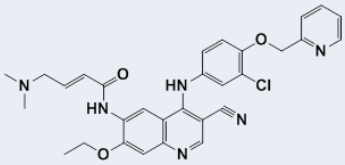
Figure 4-7: Selected fluorescent dyes for TKIs' labelling. Chemical structure, commercial name and absorption and emission maxima (Abs_{max} and Em_{max}, respectively) of selected fluorescent dyes for labelling EGFR TKIs in this study.

The last step before starting the synthesis of the fTKIs was to choose the TKIs for developing the fTKIs. In order for the potential inhibitors to be able to turn into probes for imaging, the following criteria should be met:

- the inhibitors must not have substituent on the R_4 position or at least have an ether group (methyl or ethyl);
- the structural features and the synthetic route of the inhibitor must allow the incorporation of the PEG linker and subsequently of the fluorophore;
- the inhibitors must be very selective to their primary target (or targets if they bind more than one ErbB member);
- the final list should include Type I and II as well as reversible and irreversible TKIs.

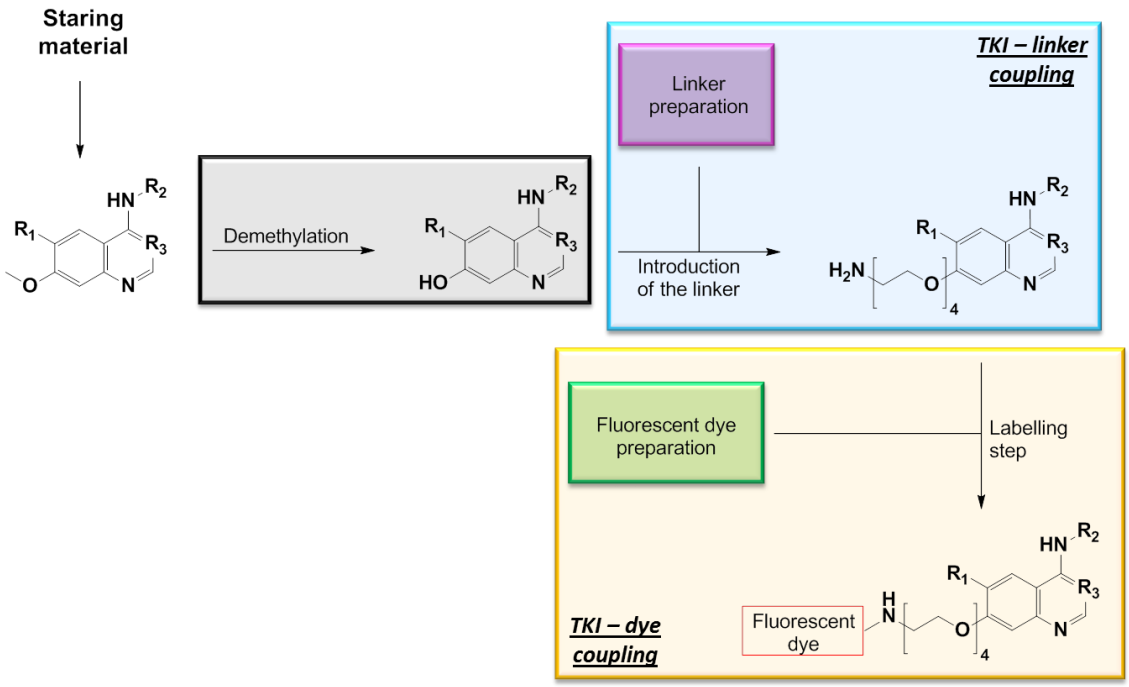
Based on the above criteria I decided upon a set of TKIs which could be potentially employed in studying all ErbB family members. They are summarized in the table below (***Table 4-1***). All of these compounds have affinities in the low nanomolar range for their targets and great selectivity[176,182]. The fact that I have chosen TKIs from all four categories (first and second generation as well as type I and II) will allow me to target all the possible conformations as well as to study the differences between them. Moreover, irreversible fTKIs should be suitable for live cell imaging.

Table 4-1: Selected inhibitors for developing fluorescent probes

Name	Structure	Protein target	Generation /Type	Status
Gefitinib (Iressa®)		EGFR	First / I	Approved
Lapatinib (Tykerb®)		EGFR/ErbB2	First / II	Approved
Pelitinib (EKB-569)		EGFR	Second / I	Discontinued
Neratinib (HKI-272)		ErbB1-ErbB4	Second / II	Clinical trials

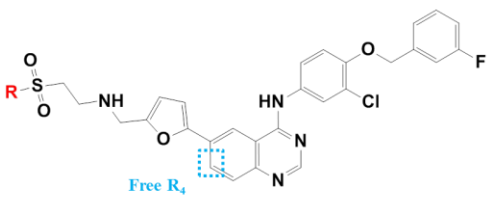
4.4 Synthesis of fTKIs

The strategy I followed for the fTKIs synthesis of the compounds bearing methoxy or ethoxy group at R_4 (Gefitinib, Pelitinib and Neratinib) was the same and is depicted in ***Scheme 4-1***. Starting from inexpensive starting materials, I prepared the anilinoquinazoline core and then demethylated the methoxy group. The free hydroxy group was then alkylated with commercial tetraethylene glycol (linker) which had previously been functionalized at the other end for later amide coupling with the dye. Synthesis of the succinimidyl ester of the dye and coupling with the TKI-linker compound at the last step afforded fTKI.



Scheme 4-1: Synthetic strategy for developing fluorescently labelled versions of Gefitinib, Pelitinib and Neratinib.

For the synthesis of fluorescent labelled Lapatinib, the strategy was slightly different due to a different labelling position compared to other three. The reason for choosing an alternative place for incorporating the linker was because Lapatinib does not have any substituent at R_4 position so introduction of any chemical group might alter specificity and/or affinity. Moreover, a structure–activity relationship (SAR) study of Lapatinib derivatives[192] revealed that substitution of the methyl sulfone on the terminal side-chain with bulkier chemical groups has only little effect on the potency (***Figure 4-8***). That gave me additional confidence that the fluorescent Lapatinib will retain the affinity and selectivity of the parent compound.



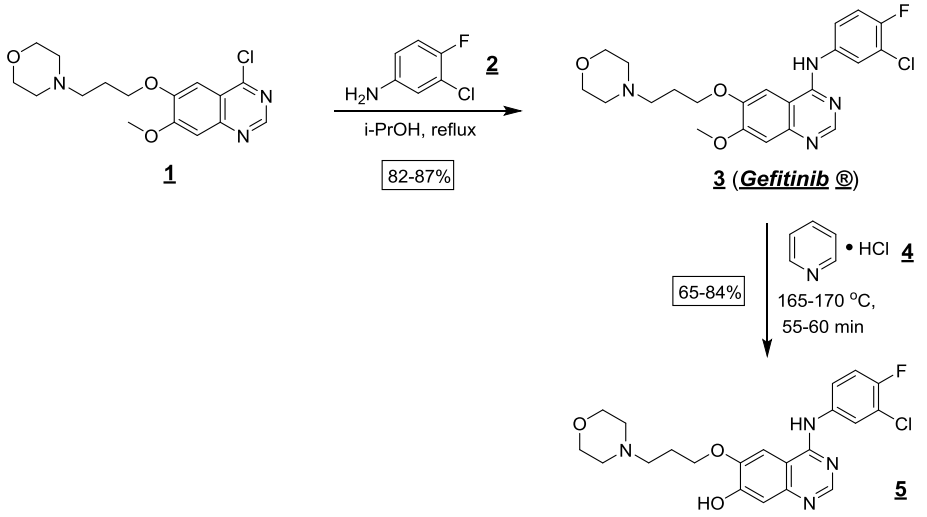
Lapatinib derivatives

R	IC_{50} (μM)				
	ErbB2	ErbB1	HN5	BT474	N87
- CH_3	0.010	0.022	0.12	0.08	0.08
n-Pr	0.026	0.021	1.6	3.14	8.66
i-Pr	0.024	0.018	0.28	0.20	0.13
- Ph	0.030	0.019	1.07	1.04	0.87
2-Pyridyl	0.030	0.019	1.55	0.97	0.79
	0.026	0.017	0.5	0.39	0.23

Figure 4-8: Sulfone derivates of Lapatinib and their inhibition and cellular efficacy (*Adapted from Petrov K. G., Bioorg. Med. Chem. Lett., 2006***)[192].**

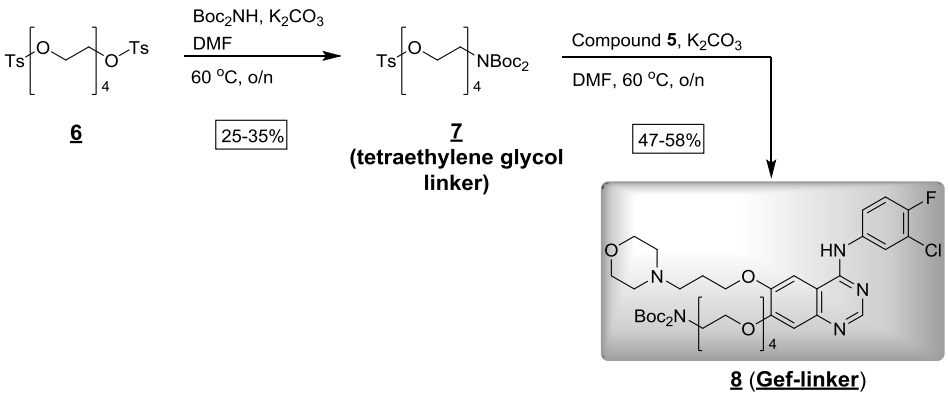
4.4.1 Preparation of modififed Gefitinib

The synthesis of Gefitinib modified with the linker began with a relative inexpensive starting material which afforded Gefitinib (**3**) in only one step after refluxing it with aniline **2**[138] (***Scheme 4-2***). For the demethylation step the procedure reported by Wang *et al.*[140], in their synthesis of radiolabelled Gefitnib as a PET agent, was followed. The demethylation reagent used was the widely used LiCl; however, in my hand the reaction gave very poor yield despite the prolonged heating. I next sought to attempt to increase the yield in a microwave-assisted reaction[141]. Even though both the reaction time and the yield were improved, the latter was still very low. Compound **5** was finally obtained with very good yield and without the need of further purification using pyridinium hydrochloride (pyHCl) in solvent-free conditions[142] (***Scheme 4-2***).



Scheme 4-2: Preparation of the demethylated Gefitinib.

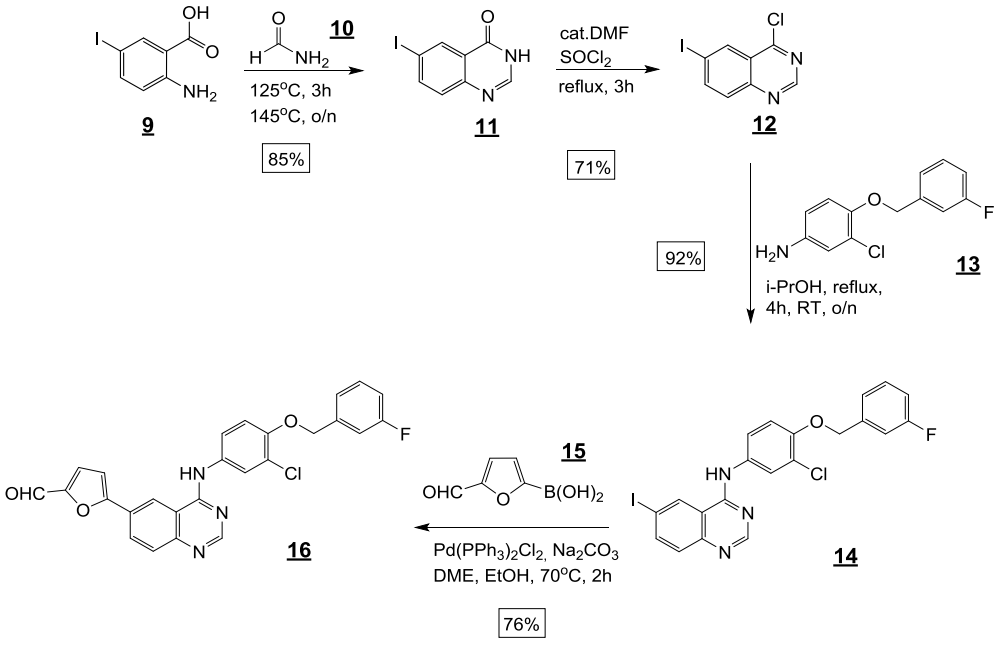
Before alkylation of the phenolic hydroxyl group of compound **5** with the linker, the one end of the linker was functionalized in order to perform the coupling with the dye at the last step. I started from tetra-ethylene glycol di-p-tosylate (**6**) and replaced one of the p-Ts esters with a protected amine as described previously[143] (***Scheme 4-3***). Finally, attachment of tetraethylene glycol linker (**7**) took place in mild conditions in an overnight reaction which furnished compound **8** (Gef-linker) in moderate yield (***Scheme 4-3***). The final step for the synthesis of fluorescently labelled Gefitinib derivatives, as well as of the rest TKIs, will be shown later (***Section 4.4.5***).



Scheme 4-3: Functionalization of tetraethylene glycol linker and coupling with compound 5.

4.4.2 Preparation of Lapatinib-linker compound

The synthesis of the equivalent Lapatinib-linker (Lap-linker) compound is shown in the next three schemes. Ring closure of 5-iodoanthranilic acid (**9**) in refluxing formamide (**10**) gave 6-iodo quinazolinone (**11**) which was subsequently chlorinated with thionyl chloride[144] (***Scheme 4-4***). Condensation of the commercially available aniline **13** with **12** afforded 6-iodo anilinoquinazoline **14** in great yield[147] (***Scheme 4-4***). The precursor (**16**) of the sulfone Lapatinib derivative was obtained by Suzuki palladium-catalysed cross coupling of **14** with 2-furanyl boronic acid (**15**) as described by Mahboobi *et al.*[145] (***Scheme 4-4***).



Scheme 4-4: Synthesis of precursor compound 16.

In order to attach the linker at the preferred position discussed earlier (***Section 4.3***) a hydroxyl group is required to participate in the reaction shown earlier for Gefitinib (***Scheme 4-2***). I therefore examined the crystal structure of Lapatinib bound to wt EGFR (***Figure 4-9***) to get an insight of the orientation of the side-chain. I noticed that the methyl group on the sulfone appears to be facing towards the protein surface, even though is poorly defined as Wood *et al.*[47] found. For that reason I decided to introduce a hydroxy group on the sulfone side-chain separated by a spacer of two carbon atoms from the sulphur atom.

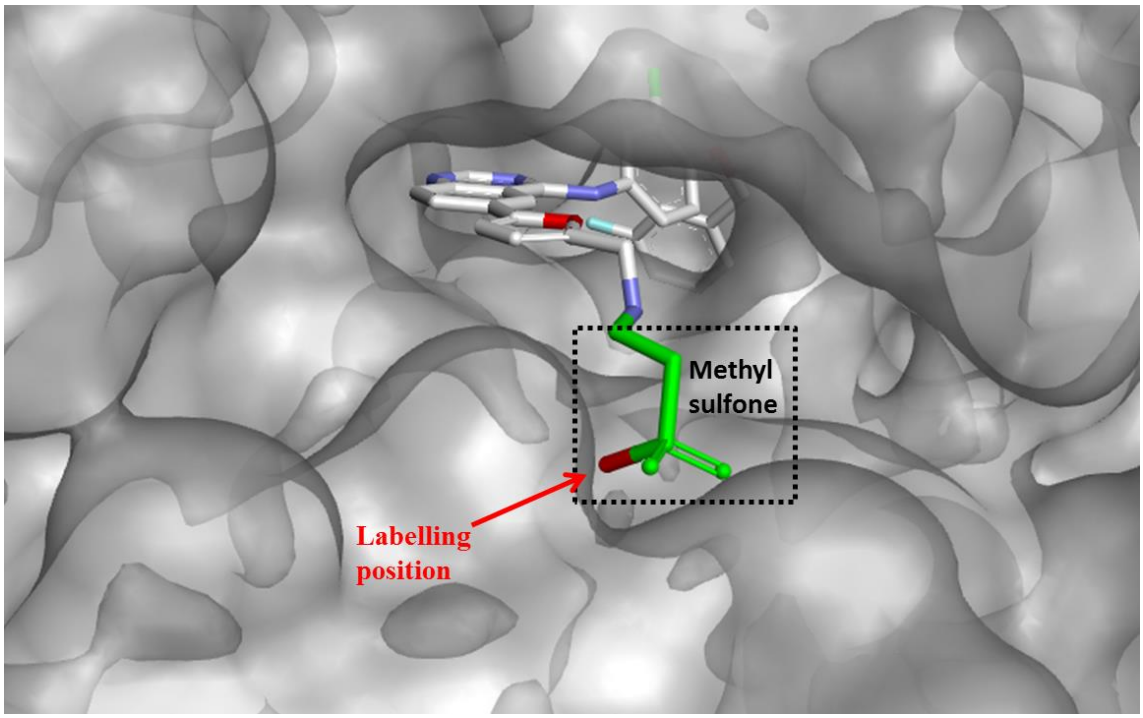
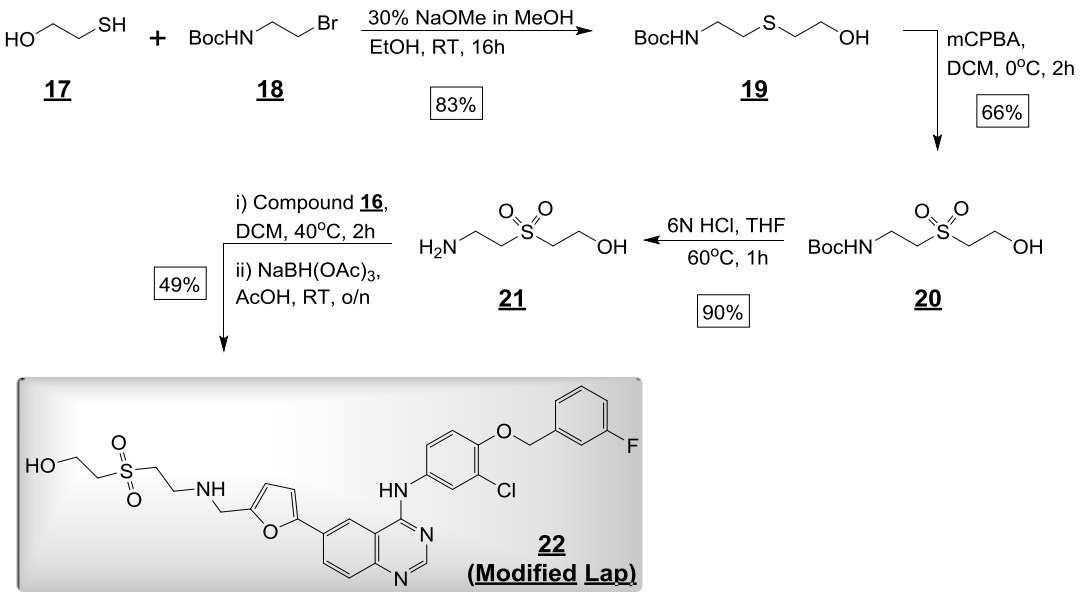


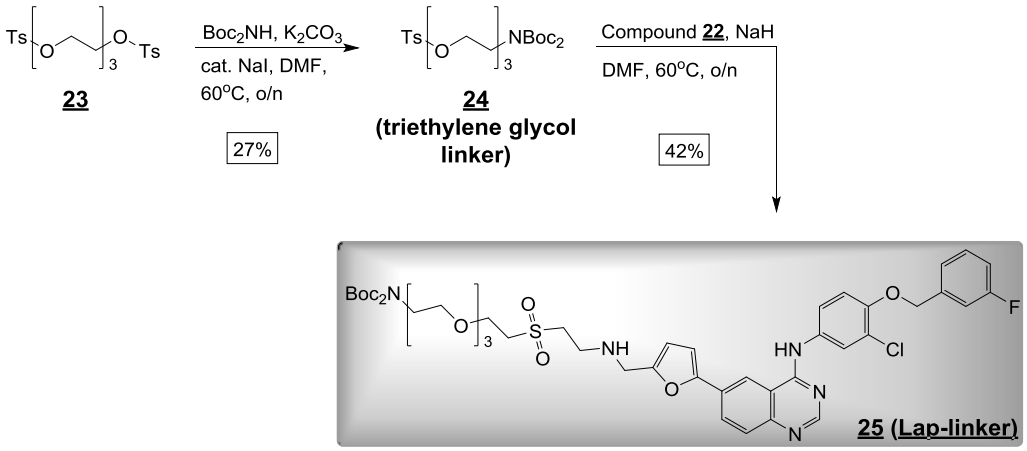
Figure 4-9: View of the labelling position on Lapatinib's side-chain (PDB: 1XKK).

The synthesis of the 2-[(2-aminoethyl)sulfonyl]ethanol (**19**, ***Scheme 4-5***) was adapted from described procedures by Seto M. and Ohashi T.[150]. The synthesis started with an S_N2 displacement of bromide from compound **18** with the thiol group of 2-mercaptoethanol, followed by oxidation of the thiol to sulfone with 3-chloroperbenzoic acid at 0^oC (***Scheme 4-5***). The amino-protective group was then cleaved before the reductive amination step which afforded the modified Lapatinib **22** (***Scheme 4-5***).



Scheme 4-5: Preparation of the sulfonyl side-chain and synthesis of the modified Lapatinib.

The selected linker for Lapatinib was slightly shorter (tri-ethylene glycol) than the linker used in Gef-linker synthesis as the sulfone side-chain was already extended away from the pocket. The preparation of the linker was performed according to the tetra-ethylene glycol synthesis (***Scheme 4-3***); however, the yield of the reaction was always lower. Even though the yield of the reaction was increased by addition of catalytic amount of NaI, it was not at the same levels as previously (compound **24**, ***Scheme 4-6***). Alkylation of compound **22** with the linker did not proceed in the same conditions as Gefitinib did. That was not unexpected as deprotonation of aliphatic hydroxy groups require stronger conditions. Therefore the potassium carbonate was replaced with a stronger base, sodium hydride, and repeated the reaction which furnished compound **25** (***Scheme 4-6***).



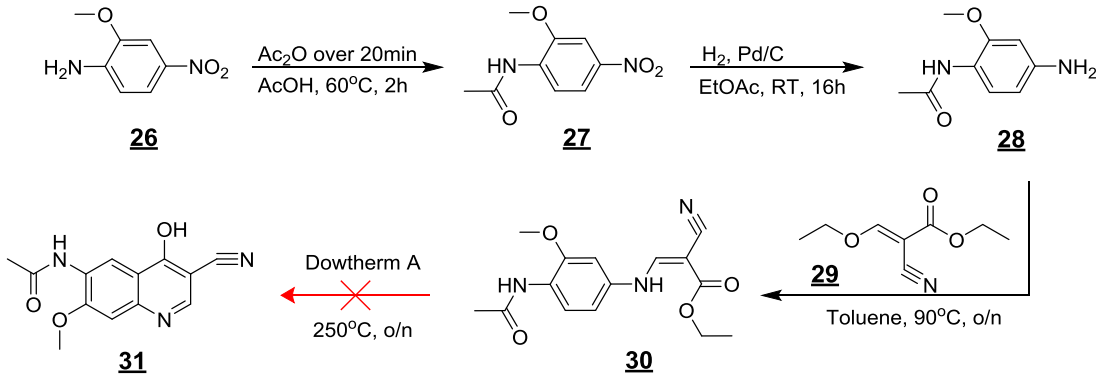
Scheme 4-6: Synthesis of Lap-linker compound.

4.4.3 Preparation of modified irreversible TKIs

Both selected irreversible inhibitors, Pelitinib and Neratinib, have been developed by the pharmaceutical company Wyeth and share the same quinoline-3-carbonitrile core (compound **40**, ***Scheme 4-9***). Their difference lies on the aniline substituent, as mentioned earlier (***Section 4.2***). This allowed me to carry out a common synthetic route up to the point of insertion of the anilines. As discussed in ***Section 4.4***, the synthetic strategy for their synthesis is the same as for Gefitnib with the sole variation in the last step which in this case is not the attachment of the linker but the incorporation of the Michael acceptor. This is due to the high reactivity of the α,β-unsaturated amides, which gives them their ability to form covalent bonds with cysteines in cell environment, and which might lead to protonation in organic reaction solutions.

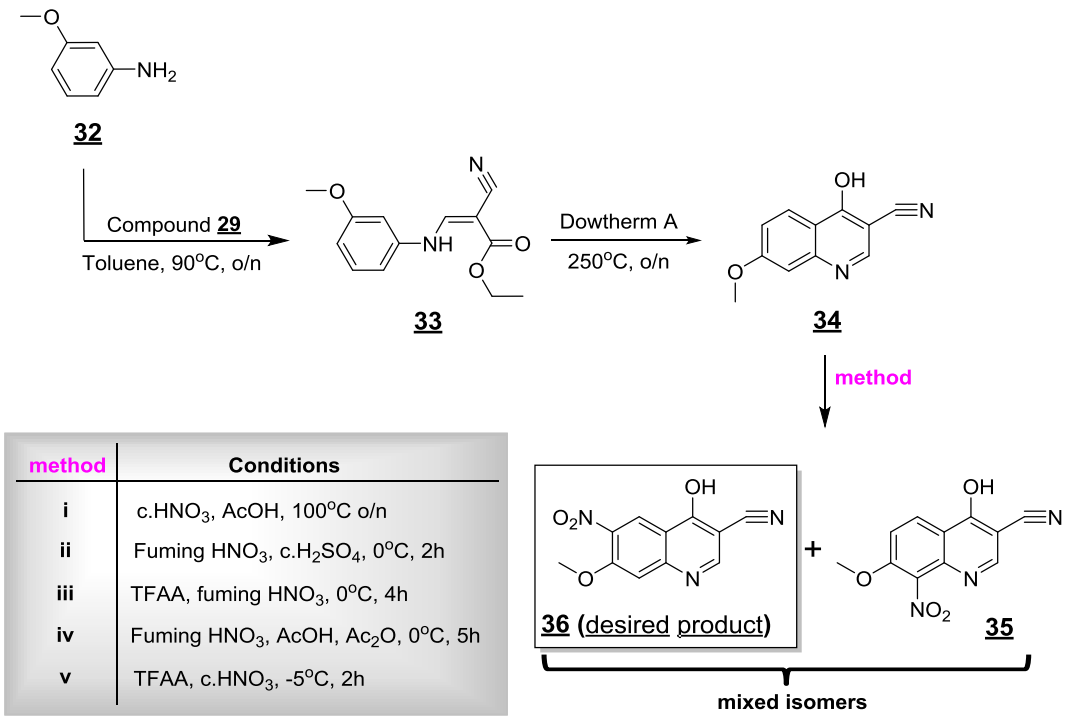
The synthesis was based on published work of the aforementioned TKIs[138,154,181,193]. The compounds were designed to carry a methoxy group instead of an ethoxy group, as the commercial one, because demethylation (step before attachment of the linker) is more efficient and faster with this group. The synthesis started from aniline **26** which was protected by acylation (compound **27**), followed by palladium-catalyzed reduction of the nitro group via hydrogenation to afford compound **28** (***Scheme 4-7***). The precursor (compound **30**) of the quinoline-3-carbonitrile (compound **31**) was synthesized by conjugate addition of the free N to the double bond of (E)-ethyl 2-cyano-3-ethoxyacrylate (compound **29**) in toluene (***Scheme 4-7***). The last step was the thermal cyclization at 250°C using a high boiling point solvent (Dowtherm A) in an overnight reaction. Despite the fact that the current step has been reported to proceed successfully,

I was unable to obtain the desired product (compound **31, *Scheme 4-7***). Judging by the state of the isolated product, it was hypothesized that polymerization may have occurred probably due to the amino-protective group.



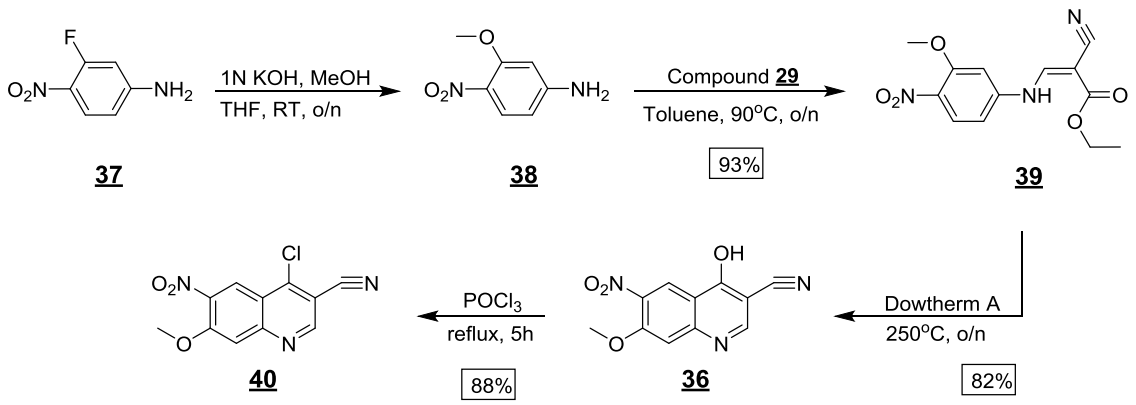
Scheme 4-7: First attempt towards the synthesis of 4-anilinoquinoline-3-carbonitrile scaffold.

Based on the results above (***Scheme 4-7***) and on my speculation, an alternative route was designed which entailed omitting the amine from the staring material and inserting it after the cyclization step. Since amines cannot directly be inserted on aromatic rings, I would first nitrate selectively on the same position, according to published procedures using exactly the same substrate[138], and then reduce the nitro group. As hypothesized, thermal cyclization took place successfully and with an excellent yield, affording compound **34** (***Scheme 4-8***). Although the nitration (**method i**, table in ***Scheme 4-8***) reaction was allowed to proceed to completion (no reactant detected by LC-MS), both ^{1}H NMR and HPLC analysis revealed the presence of two nitro products (compounds **35** and **36**, ***Scheme 4-8***). The same result, with variation on the ratio of the isomers, was obtained with all the possible nitration conditions tested (**methods ii-v**, table in ***Scheme 4-8***). These isomers (compounds **35** and **36**) have identical properties and therefore purification and isolation of the desired product at that stage would be impossible. Separation at later stage of the synthesis might have been possible however at least half of the final product would have been useless.



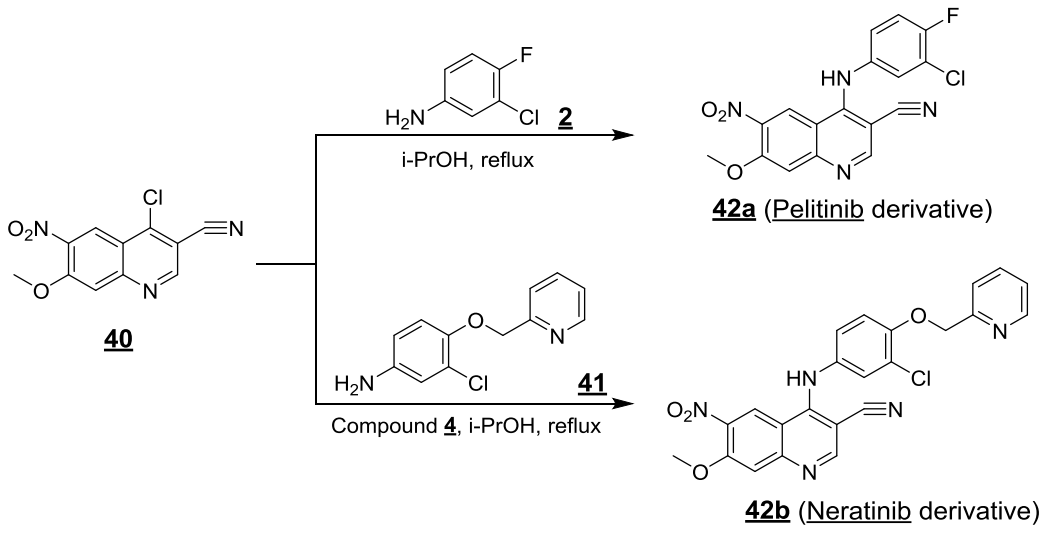
Scheme 4-8: Second attempt towards the synthesis of 4-anilinoquinoline-3-carbonitrile scaffold.

Taking into consideration the issues I encountered with both the aforementioned strategies (***Scheme 4-7*** and ***Scheme 4-8***), an alternative synthetic route was designed where the amine would not be present before cyclization and the nitro group would be on the starting material at the desired position. Due to commercial availability reasons, I started from 3-fluoro-4-nitroaniline (compound **37**) which was converted to the methoxy-substituted compound **38** after aromatic substitution of the fluorine[151] (***Scheme 4-9***). Conjugate addition to compound **29** and thermal cyclization, both steps performed as in ***Scheme 4-8***, afforded the desired nitro-compound **36** (***Scheme 4-9***). In order to create an attachment point for the subsequent addition of the anilines that would result into Pelitinib and Neratinib derivatives, the nitro-compound **36** was converted into the aromatic chloride **40** using phosphoryl chloride (***Scheme 4-9***).



Scheme 4-9: Third attempet and synthesis of irreversible fTKIs' scaffold.

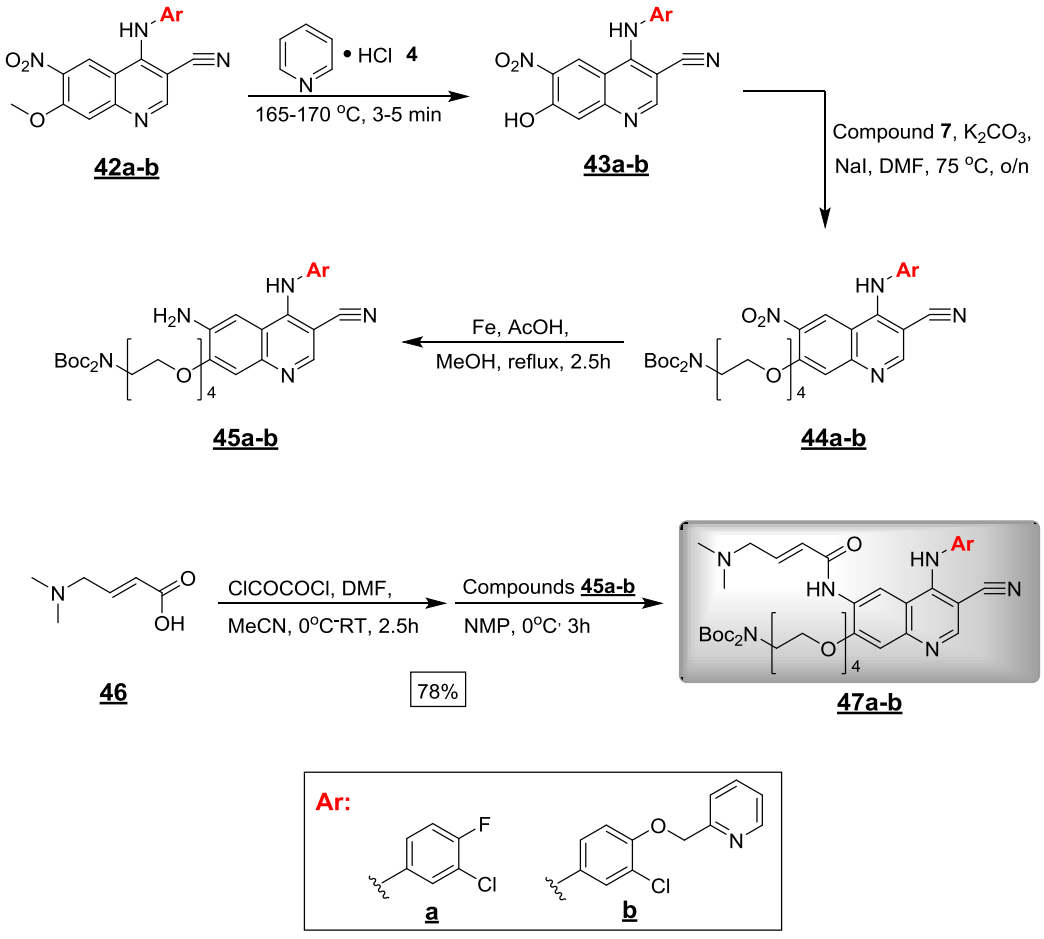
The following step was the aromatic substitution of the chloride of compound **40** by either aniline **2**[138] (earlier used in Gefitinib synthesis, ***Scheme 4-2***) or aniline **41**[181] which afforded Pelitinib (**42a**) and Neratinib (**42b**) derivatives, respectively (***Scheme 4-10***).



Scheme 4-10: Condensation of aniline 2 or 41 for synthesis of Pelitinib and Neratinib derivatives.

The remaining synthetic steps towards the irreversible TKI-linker compounds were the same for both Pelitinib and Neratinib as shown in ***Scheme 4-11*** and they were performed in parallel reactions. The first step was demethylation, carried out using the same condition described for Gefitinib (**Scheme 4-2**). In contrast to demethylation of Gefitinib, the reaction on these substrates was extremely fast and no starting material was detected after only 3-5 min. The conditions for introducing tetraethylene glycol

linker **7** (***Scheme 4-3***) were also different; catalytic amount of NaI and higher temperature (75°C instead of 60°C) were required to push the reaction to completion (compounds **44a-b**). Reduction of the nitro group with iron metal resulted in the aniline (compounds **45a-b**) which was then coupled with the acyl chloride of **46** (which had been previously formed *in situ* with oxalyl chloride) to afford the final products **47a** (Pel-linker) and **47b** (Ner-linker)

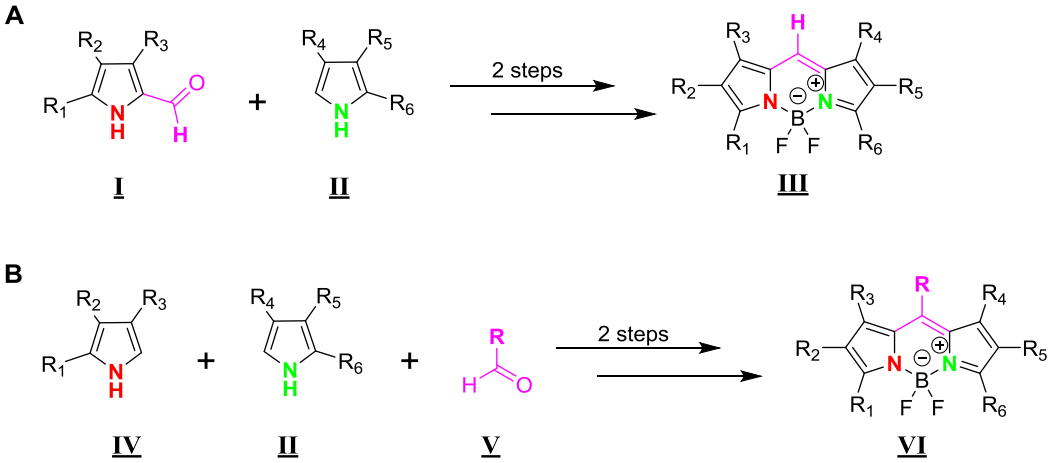


Scheme 4-11: Synthesis of irreversible Pelitinib- (47a) and Neratinib-linker (47b) compounds

4.4.4 Synthesis of fluorescent dyes

Synthesis of the Bodipy dye core (boron-dipyrromethene) can be performed in two steps either in a 2- or in a 3-componet reaction[194] (***Scheme 4-12***). The former route (***Scheme 4-12A***) requires an 2-acylpyrrole (**I**) and a pyrrole with an unsubstituted α-carbon (**II**). The latter route (***Scheme 4-12B***) on the other hand requires two α unsubstituted carbon and an aldehyde (or acyl chloride) and it is preferred for the preparation of the so-called *meso*-substituted Bodipy dyes (**VI**). Substitution on the

pyrrole rings can be performed before or after the formation of the boron-dipyrromethene core (**III**), depending on the chemistry allowed on them. Substituents that extend the π-conjugated system (electron delocalization) shift the emission from low to high wavelength. This is applicable to all types of dyes.



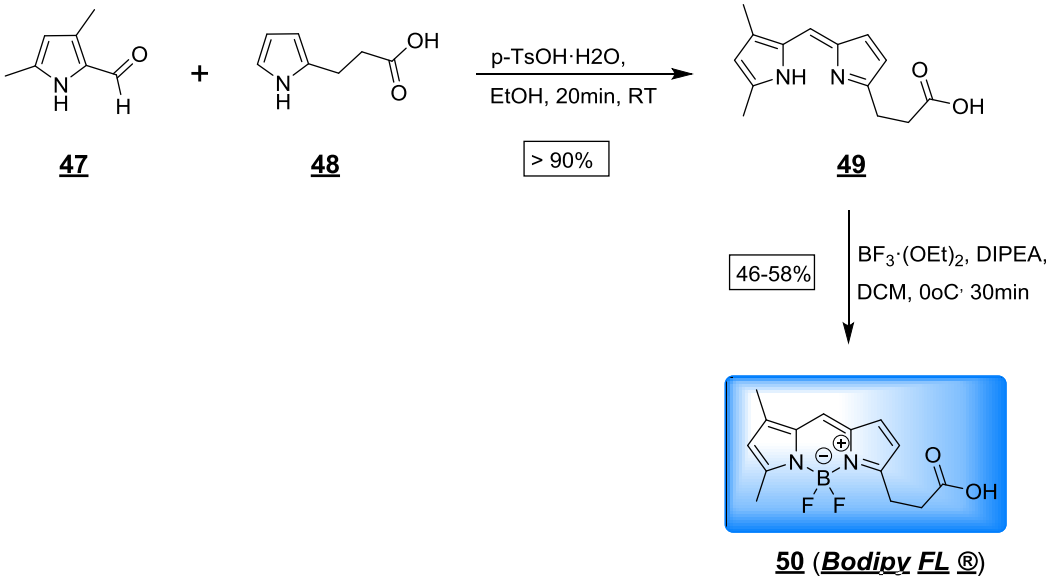
Scheme 4-12: 2- and 3-component strategies for synthesis of Bodipy dyes

In order to be able to conjugate the dyes (both Bodipy and TMR) I require a free carboxylic acid group which would be activated by forming the succinimidyl ester and then coupled with the TKI-linker compounds via an amide bond. The synthesis of the commercial Bodipy FL® (Bod FL) has been previously reported in two studies[156,195] (in addition to the patent from Molecular Probes®[196]). However, in both studies the propanoate group had to be protected at the beginning of the reaction, either by esterification or by a trichloroethyl group, and deprotected at the end. As a result, Bod FL synthesis requires two more steps than one would expect, based on the general scheme above, and also the overall yield of the synthesis drops significantly.

I attempted an alternative synthesis based on a published procedure by Meltola *et al.*[195] where the authors synthesized a Bodipy dye with the propanoate group but in a different position. Even though the formation of the desired product was indeed detected, the yield of the reaction was really low. Therefore, this procedure was abandoned.

The solution was found in a Thesis by Spicka K. J[155] (Montana State University). He managed to eliminate the protection-deprotection step by using *para*-toluenesulfonic acid (p-TsOH) which allowed the reaction to proceed with good yield. The fact that the

reaction was successfully applied to a variety of different substrates enhanced my confidence. Addition of p-TsOH in the reaction mixture of compound **47** and **48** instantaneously changed the colour of the solution, indicative of product formation (***Scheme 4-13***). The dipyrromethene **49** was generated in excellent yield (almost full conversion to product) and with high kinetic rates. Transformation to the boron complex (**50**) was then performed as reported by Gießler *et al*[156] with a moderate yield.

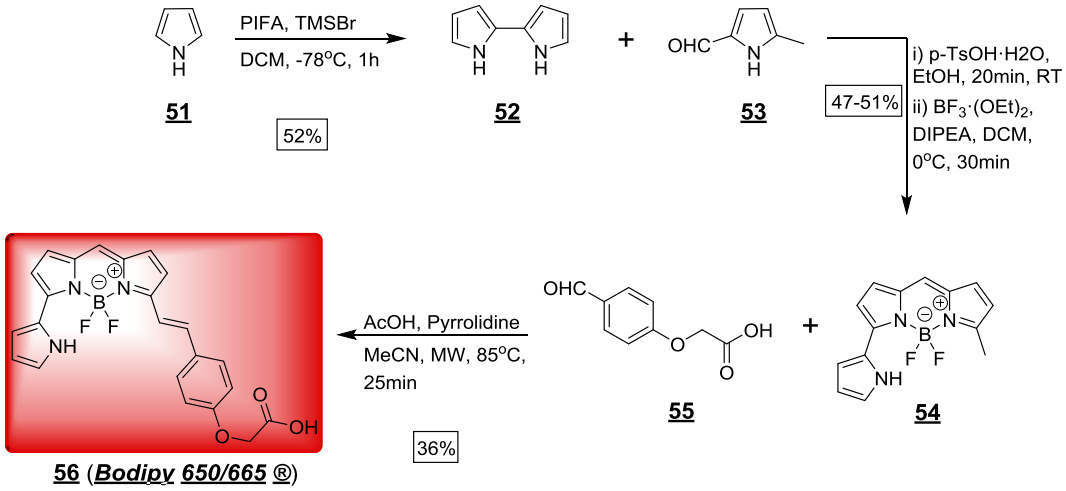


Scheme 4-13: Synthesis of Bodipy FL

The far-red emission of Bodipy 650/665 (Bod 650) is due to substituents on both the right- and the left-hand side pyrroles which extend electron delocalization. On the right-hand side, there is an α-substitution from another pyrrole ring whereas on the left an α-substitution from a styryl group. Pyrrole rings can be introduced on the final molecule but only on *meso*-substituted Bodipy dyes[197] (***Scheme 4-12B***, **VI**), therefore the bipyrrole **52** was decided to be prepared beforehand. Among the few procedures found in literature for synthesising bipyrroles, the one reported by Dohi *et al.*[157] not only overcomes the issues from which the others suffer, but also proceeds with good product yields. On the other hand, styryl substituents can be incorporated before[198] and after[158] cyclization, however the former procedure requires two additional steps.

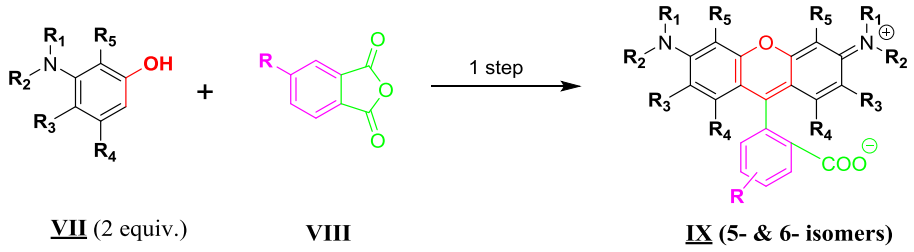
The synthesis of Bod 650 (***Scheme 4-14***) started with the oxidative coupling of the pyrrole **51**, which generated α-linker bipyrrole **52**, followed by formation of Bodipy core (compound **54**) as described for Bod FL earlier (***Scheme 4-13***). Finally, a

microwave-assisted Knoevenagel-type condensation of the acid **55** with Bodipy **54** furnished the desired product (compound **56**, ***Scheme 4-14***).



Scheme 4-14: Synthesis of the far-red Bodipy dye.

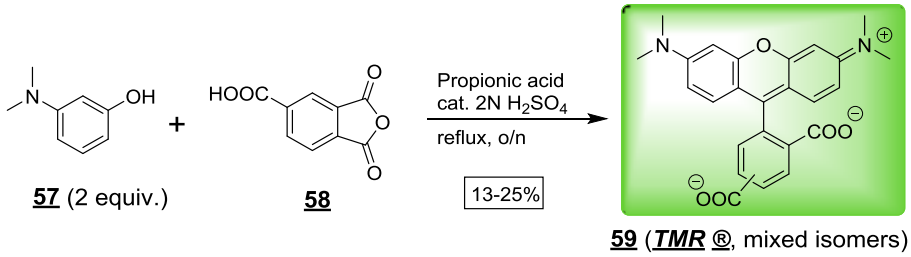
The last fluorescent dye to be synthesized is the tetramethyl carboxyrhodamine. The most common approach for the synthesis of rhodamine-based (and fluorescein-based) dyes[199] is between 2 equivalents of aminophenol (**VII**) and 1 equivalent of phthalic anhydride (**VIII**) in the presence of a Lewis or a protic acid at high temperature (***Scheme 4-15***). The reaction is not regioselective and therefore both 5- and 6- isomer of TMR are generated in the mixture. The two isomers can be separated by preparative HPLC or chemical separation[160].



Scheme 4-15: General synthesis of Rhodamine-based dyes

Rhodamine-based dyes, being among the oldest dyes reported[200], have been well explored in order to improve the synthetic efficiency but also to eliminate the use of harsh conditions such as concentrated sulphuric acid at 170°C. Among the various conditions tested[160,201,202] the procedure reported by Uddin et al.[159] gave me the best results in terms of product yield and ease of purification and isolation. The reaction is

shown in ***Scheme 4-16*** and took place in one step by condensation of 3-(dimethylamino) phenol (compound **58**) and phthalic anhydride (compound **58**) in a weakly acidic solvent in the presence of catalytic amount of sulphuric acid.

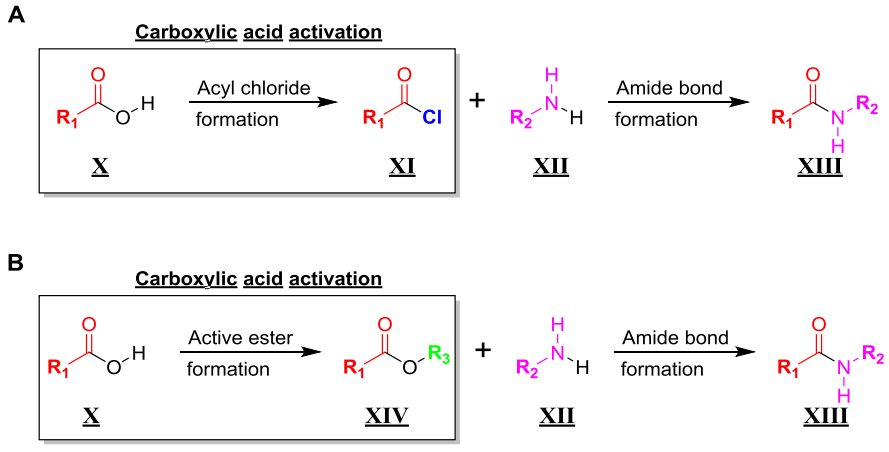


57 (2 equiv.) **58** **59** (***TMR*** ®, mixed isomers)

Scheme 4-16: Synthesis of tetramethyl 5/6-carboxy-Rhodamine

4.4.5 Coupling of TKI-Linker Compounds with Dyes

The last part of the synthesis is the coupling of the TKI-linker compounds (compounds **8**, **25**, **47a** and **47b**) with the fluorescent dyes (compounds **50**, **56** and **59**) through formation of an amide bond. This type of reaction is facilitated by activation of the carboxylic acid group before reacting with the amine[203]. The two most widely used methods for activation of carboxylic acids are by formation of acyl chloride (***Scheme 4-17A***) or by formation of active esters (***Scheme 4-17B***) such as succinimidyl (NHS) esters. The latter is preferred for the protection of the dyes due to its non-acidic conditions.

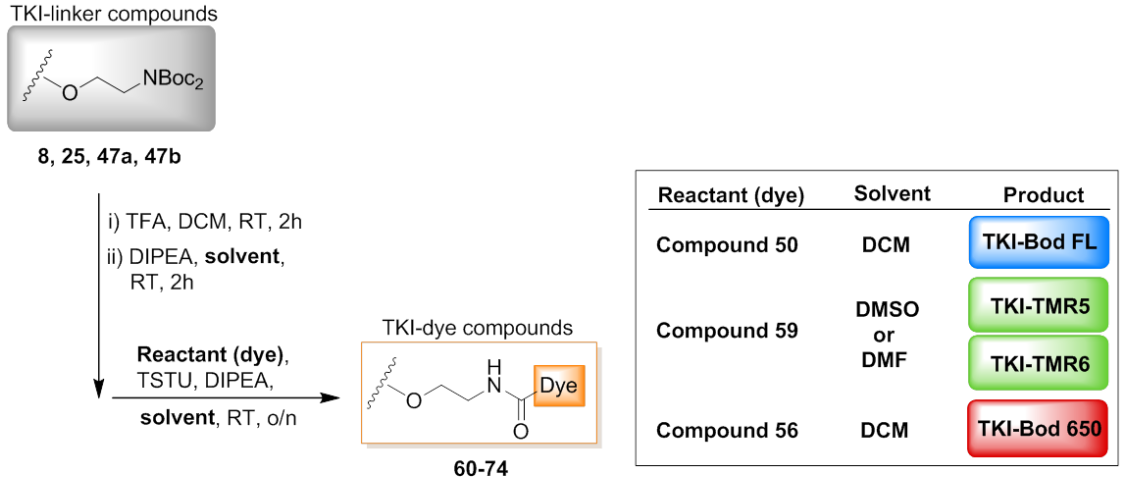


Scheme 4-17: Most commonly used strategies for amide bond formation.

In this synthesis the activated esters were generated and reacted with the amine *in situ* eliminating the need for extra purification and isolation steps[159]. Moreover in this

fashion, the overall yield starting from the free carboxylic acid to amide (2 steps) was higher as attempts to purify the NHS esters resulted in transesterification of the dye with the methanol (which was required for solubilizing the dyes).

The method used for the TKI-dye coupling was the same for all TKIs and all dyes except for the solvent which was chosen based on the solubility of the dye (table in ***Scheme 4-18***). The two amino-protecting groups of the TKI-linker compounds (**8**, **25**, **47a** and **47b**) were first cleaved using trifluoroacetic acid (TFA) followed by removal of unreacted TFA. The residue was then re-dissolved and kept in a separate container (c1). In another container (c2), the "free" dye was left to be activated for 2 h using tetramethyl-O-(N-succinimidyl) uronium tetrafluoroborate (TSTU) before it was added to c1 to generate the fluorescent TKIs (compounds **60**-**74**). This reaction was repeated for all the possible combinations between a TKI and a dye (16 reactions), resulting in 15 TKI-dye products which were purified by preparative HPLC.



Inhibitor	Fluorescent dye			
	Bodiby FL	TMR 5	TMR 6	Bodipy 650
Gefitinib	Compound **60**	Compound **64**	Compound **67**	Compound **71**
Lapatinib	Compound **61**	Compound **65**	Compound **68**	Compound **72**
Pelitinib	Compound **62**	---	Compound **69**	Compound **73**
Neratinib	Compound **63**	Compound **66**	Compound **70**	Compound **74**

Scheme 4-18: Inhibitor – dye coupling.

In this chapter, the design and synthesis of fluorescently labelled ErbBs' inhibitors was shown. With regards to the selection of the TKIs, an extensive literature review for understanding and identifying highly specific inhibitors led to the selection of four compounds: Gefitinib, Lapatinib, Pelitinib and Neratinib. The fact these compounds belong to different classes (type I or II/reversible or irreversible) could potentially allow me to get an insight into their mechanism of action.

The insertion of the linker (and consequently of the dye) was rationally design in such a way in order not to alter (or at least to have the minimum effect) the binding of the compounds. Molecular modelling was also performed in order to predict the optimum linker length.

Finally, the synthesis of every compound reported herein (including the dyes) was based on previously reported synthesis. Even though some of the literature procedures did not work as expected, alternative reactions led to the desirable final compounds. Overall 15 final fluorescent EGFR inhibitors were synthesised.

5 Evaluation of fTKIs as Imaging Probes and their Applications in Imaging Techniques

5.1 Introduction

The synthesized fluorescent ErbBs inhibitors described in ***Chapter 4*** are shown in ***Figure 5-1***. The final "palette" includes TKIs from 4 different categories tagged with three different fluorescent dyes each.

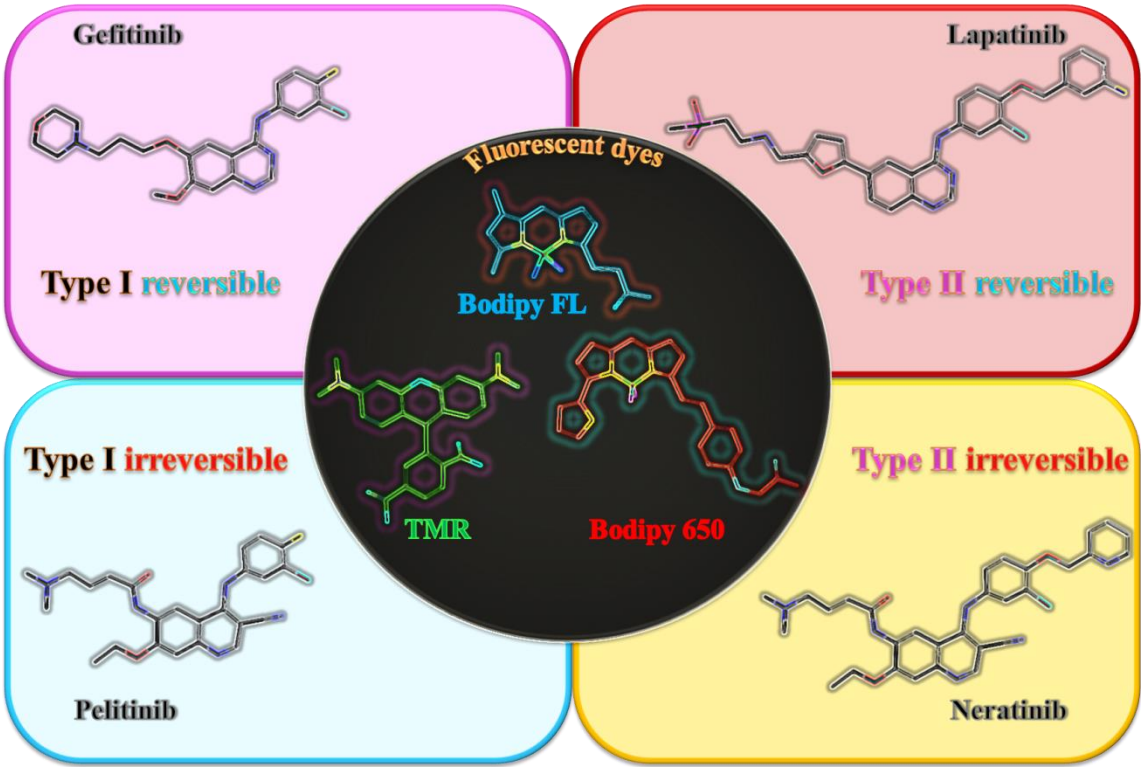


Figure 5-1: Overview of the synthesized fluorescent ErbB inhibitors

In order to be able to use the fTKIs as probes for imaging it is critical to evaluate if, and to what extent, the conjugation of the linker and the dyes has affected properties essential for their function. These properties are:

i. Binding affinities against the ErbB family members
ii. Fluorescent properties
iii. Cell permeability
iv. Binding/labelling in cells

v. Binding specificity

Following this evaluation, I will explore the suitability for using fTKIs in various imaging techniques and described the development of methods that would allow me to probe the target receptors and to study the mechanism of action of TKIs.

5.2 Characterization of fTKIs

The binding affinity of any novel imaging probe (from small-molecule to antibodies) is surely the foremost property and could define the 'fate' of the probe. Low affinity probes would require high concentration for engaging the target which would affect the signal-to-noise ratio or might result in high non-specific binding.

As discussed earlier, the TKIs were selected due to their high specificity to the EGFR family. The selectivity profile data I reviewed in the selection process were obtained from previous studies investigating the selectivity of various inhibitors against the kinome[176,182,183]. The platform used in these studies for screening the compounds (KinomeScan®, initially developed and reported by Fabian *et al.*[183]) is offered by DiscoveRx (San Diego, USA), so it was decided to determine the binding affinities of the fTKIs using the same assay. An overview of the assay's principle is shown below.

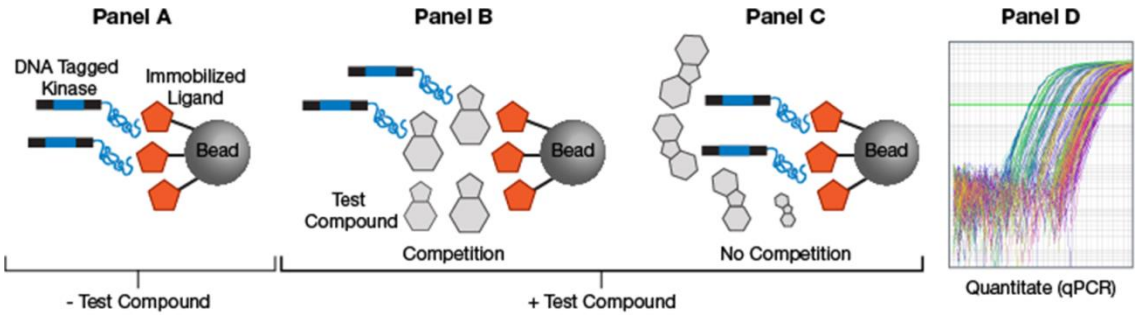


Figure 5-2: KinomeScan® assay for determination of K_d values of fTKIs (*Adapted from DiscoveRx website***).**

Binding affinities of fluorescently labelled Gefitinib (f-Gef) and Pelitinib (f-Pel) were only determined for EGFR (since there are selective to EGFR), whereas for the multi-targeted Lapatinib (f-Lap) and Neratinib (f-Ner) affinity for the other members of the ErbB family was also assessed (***Table 5-1***). Furthermore, in order to ascertain that the dyes of the conjugates have not contributed to the observed affinity, the dissociation constants (K_ds) of the modified compounds prior to dye coupling (TKI-linker

compounds) was determined. If the dye of a TKI-linker-dye compound binds to the target, the K_d of the TKI-linker compound should be higher. However, if the dye does not bind, the K_d of the inhibitor without the dye is expected to be lower as attachment of the bulky dye should at least slightly affect the affinity of the conjugate.

Table 5-1: Binding affinities of fTKIs and crucial intermediates.

	K_d (nM) *			
Compound	**EGFR**	**ErbB2**	**ErbB3**	**ErbB4**
Gefitinib® **	0.62	3500	790	410
Gef-linker (**8**)	4.9	-	-	-
Gef-Bod FL	13	-	-	-
Gef-TMR 5	39	-	-	-
Gef-TMR 6	12	-	-	-
Gef-Bod 650	29	-	-	-
Lapatinib® **	2.4	7	5500	54
Modified Lap (**22**)	3.3	8.4	-	38
Lap-Linker (**25**)	8.9	39	-	290
Lap-Bod FL	7.9	13	-	140
Lap-TMR 5	14	8.2	-	230
Lap-TMR 6	45	38	-	1500
Lap-Bod 650	18	18	-	340
Pelitinib® ***	0.44	500	-	21
Pel-Linker (**47a**)	5.1	-	-	-
Pel-Bod FL	9.3	-	-	-
Pel-TMR 6	5	-	-	-
Pel-Bod 650	28	-	-	-
Neratinib® **	1.1	6.0	7.7	2.4
Ner-Linker (**47b**)	2.9	13	25	8.2
Ner-Bod FL	4	7.6	23	17
Ner-TMR 5	7.5	6.6	12	17
Ner-TMR 6	8.1	14	18	19
Ner-Bod 650	10	33	58	17

* *Kd determination (in duplicates) was performed in DiscoveRx, for all compounds, except where noted*

** *Obtained from Davis et. al., Nat Biotechnol. 2011*[176]

*** *Obtained from Karaman et. al., Nat Biotechnol. 2008*[182]

As reported in ***Table 5-1***, all fTKIs retained low-nanomolar affinities against their targets, therefore vindicating the design of the probes. Moreover, the selectivity profile of the multi-targeted fTKIs against ErbB family members exhibited the same trend as the parent compounds (orange coloured rows). Comparison of the K_d of the TKI-linker compounds with their labelled (with any dye) versions clearly indicates that the dyes do not interact with the targets. The data also suggests that the affinity of the conjugates is affected by the size of the dye. Insertion of the Bod FL, which is 140 Da (***Table 5-2***) lighter than the other two, had the smallest effect on the affinity whereas Bod 650 (only 3 Da heavier than TMR, but with a larger cross-section) conjugates exhibited the highest K_d values.

Table 5-2: Characteristic of fluorescent dyes

Fluorescent dye	Molecular Weight (Da)	Net charge	CLogP*
Bod FL	305.1307	0	2.868
TMR	443.4944	0	-4.467
Bod 650	446.2569	0	4.979

* Hydrophilicity (ClogP) of the dyes was calculated from their structures using ChemDraw software (CambridgeSoft®)

Overall, the binding affinity data show that the fTKIs have the potential to be used for labelling all four ErbB members at the same time by using a f-Ner compound. Furthermore, it might be possible to visualize individual members by employing unlabelled and labelled compounds using one of the following strategies:

a. Either f-Gef or f-Pel for EGFR labelling
b. f-Lap with unlabelled Gefitinib (or other type I TKIs) for ErbB2 specific labelling
c. f-Ner with unlabelled Afatinib (binds to EGFR, ErbB2 and ErbB4) for ErbB3 labelling
d. f-Ner with combination of unlabelled Lapatinib (EGFR and ErbB2) and Bosutinib (ErbB3) for labelling of ErbB4.

The next step in the evaluation process was to determine whether the conjugation has altered the fluorescent properties of the dyes. For that the excitation spectra of the fTKIs were recorded and compared with the ones from the ‘free’ dyes (obtained from the

Invitrogen® website). In all three cases, the observed excitation of the fTKIs was slightly shifted (due to conjugation) towards longer wavelengths compared to commercial unconjugated dyes but without any significant change on the shape of the individual spectra.

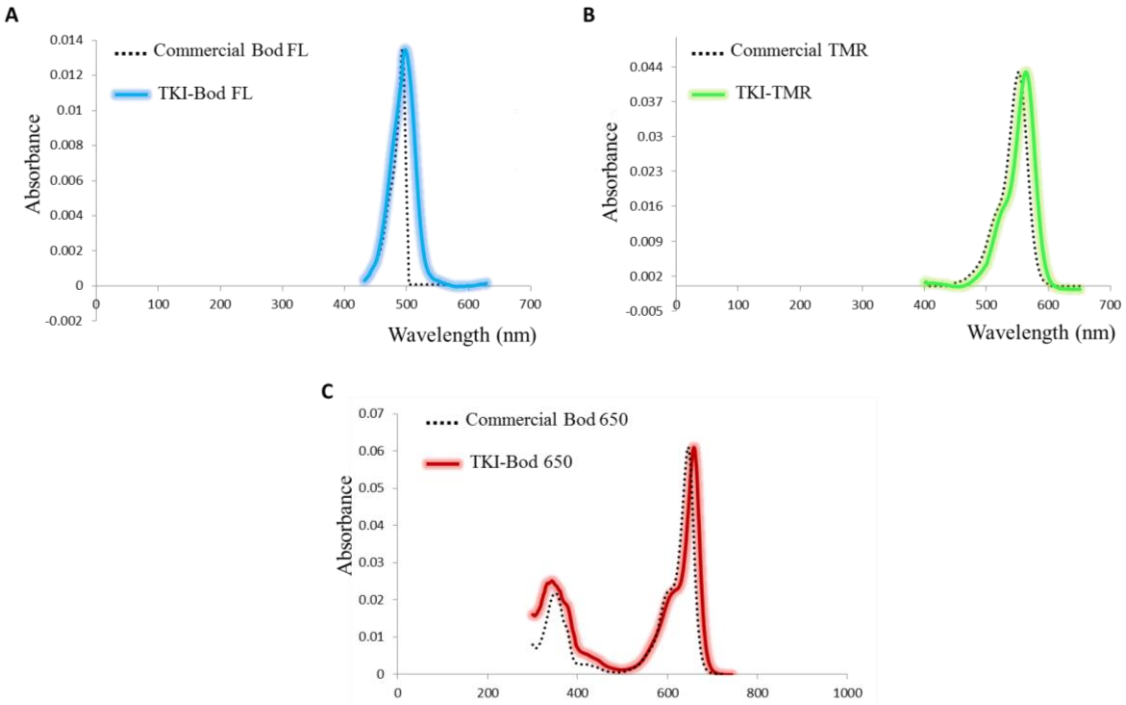


Figure 5-3: Comparison of excitation spectra of commercial "free" dyes and dyes conjugated to TKIs. Excitation spectra of Bodipy FL (**A**), TMR (**B**) and Bodipy 650 (**C**) conjugates and comparison with spectra from the free dyes (dotted line), obtained by Invitrogen.

Next, the ability of the fluorescent compounds to permeate the plasma membrane was evaluated. The fact that the inhibitors are fluorescent allowed me to assess their permeability in real-time (live-cell imaging) rather using assays such as Caco-2 permeability assay. Confocal microscopy allows optical sectioning of thick samples and therefore visualization of intracellular distribution is possible. The experiments were performed in live Chinese Hamster Ovary (CHO) cells which do not express appreciable levels of EGFR family receptors. CHO cells seeded on glass-bottomed dishes were placed on an incubator-equipped microscope pre-warmed at 37^{o}C and treated with various concentrations of fTKIs. Images were then recorded for 20-30 min using the appropriate laser beam for the dye used.

Table 5-3: Permeability of fTKIs

	Permeable	
Compound	**Yes**	**No**
Gef-Bod FL	√	-
Gef-TMR	-	√
Gef-Bod 650	√	-
Lap-Bod FL	√	-
Lap-TMR	√	-
Lap-Bod 650	√	-
Pel-Bod FL	√	-
Pel-TMR	-	√
Pel-Bod 650	√	-
Ner-Bod FL	√	-
Ner-TMR	√	-
Ner-Bod 650	√	-

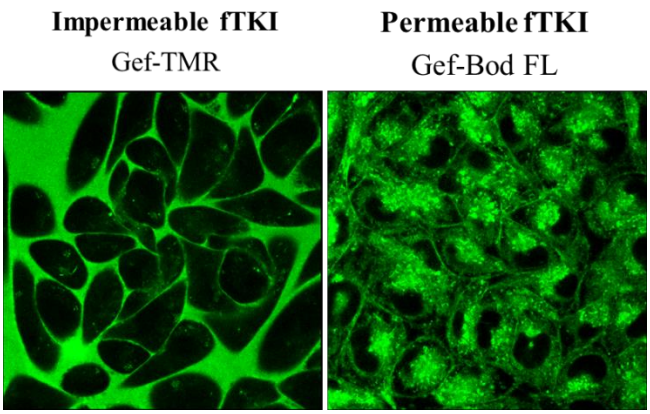


Figure 5-4: Example of an impermeable and a permeable fTKI. Live CHO cells were treated with fluorescent inhibitors *in situ* and images were recorded for 20 min at 37°C. An example of an impermeable fTKI (Gef-TMR) is shown on the left and one of a permeable (Gef-Bod FL) is shown on the right.

Out of the 12 different TKI-dye combinations only two fTKIs were not permeable: TMR labelled Gefitinib and Pelitinib (red highlighted rows, ***Table 5-3***). Surprisingly, TMR did not affect the permeability of Lapatinib and Neratinib conjugates. The difference between these two sets of conjugates is that both Lapatinib and Neratinib are ~100 Da heavier than the other two TKIs and the TMR dye. Taking also into account the hydrophilicity of TMR (CLogP: - 4.467, ***Table 5-2***), the permeability results suggest that the size of the warhead might be responsible for overcoming the resistance of a slightly hydrophilic neutrally charged dye to cross the plasma membrane. These considerations can provide an additional guidance for low-MW weight probe development in the future. It should be noted that even though Gef-TMR and Pel-TMR were not permeable they were used in imaging experiments in pre-fixed cells (see below).

Having assessed the permeability of the compounds, next I investigated the binding/labelling *in vitro* for the permeable ones. For this experiment, CHO cells stably transfected with wt EGFR expressing 400.000 receptors per cell were used. Since binding of TKIs is a relatively slow process due to their constant competition with ATP, the fTKI incubation time was increased to 2 h in order to maximize the chances of effective labelling. The receptor was also tagged on its ECD with fluorescently labelled ligands (either anti-EGFR Affibody (Affy), which does not activate the receptor[70], or

EGF) which are known to be specific for EGFR in order to examine their colocalization with the fTKIs. After labelling, cells were washed three times to remove unbound molecules, fixed with 3% paraformaldehyde (PFA) and 0.5% glutaraldehyde (GA) to avoid receptor internalization and imaged with a confocal microscope at room temperature (RT).

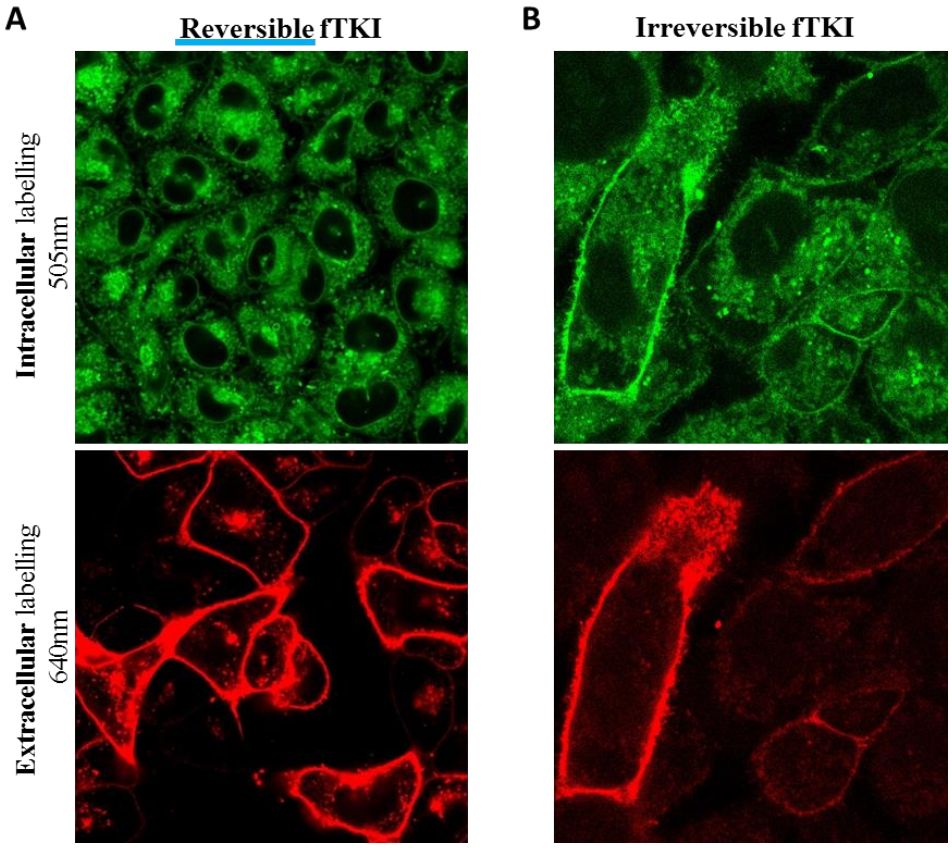


Figure 5-5: *In vitro* binding of reversible and irreversible fTKIs in live cells. Live CHO cells expressing wt EGFR were pre-treated with fTKIs for 2 h, washed thrice and fixed before imaging. Examples of labelling with a reversible (**A**) and an irreversible (**B**) fTKI are shown. Top images show intracellular labelling with fTKIs and the bottom ones extracellular labelling with EGFR Affibody.

There was not significant membrane labelling from the reversible fTKIs, despite the long incubation and high concentration (***Figure 5-5A***). This, however, was somehow expected as even bound molecules would be washed out due to their reversible binding. The membrane signal on cells pre-treated with irreversible fKTIs on the other hand was clearly present, proportional to ECD labelling and with excellent colocalization (***Figure 5-5B***). Therefore, irreversible fTKIs can be used for live cells imaging.

In parallel to the previous method, reversible fTKI labelling was also attempted to be visualized in live cells without removing the fTKI media after the 2 h treatment.

However no safe conclusion could be drawn as the presence of the fTKI in the solution made it impossible to distinguish between potentially bound molecules and molecules on the periphery of non-expressing cells (***Figure 5-6***). Upon removal of the fTKI media, without further washings, the signal on the cell edges immediately dropped to similar levels as seen in fixed cells in both expressing and non-expressing cells.

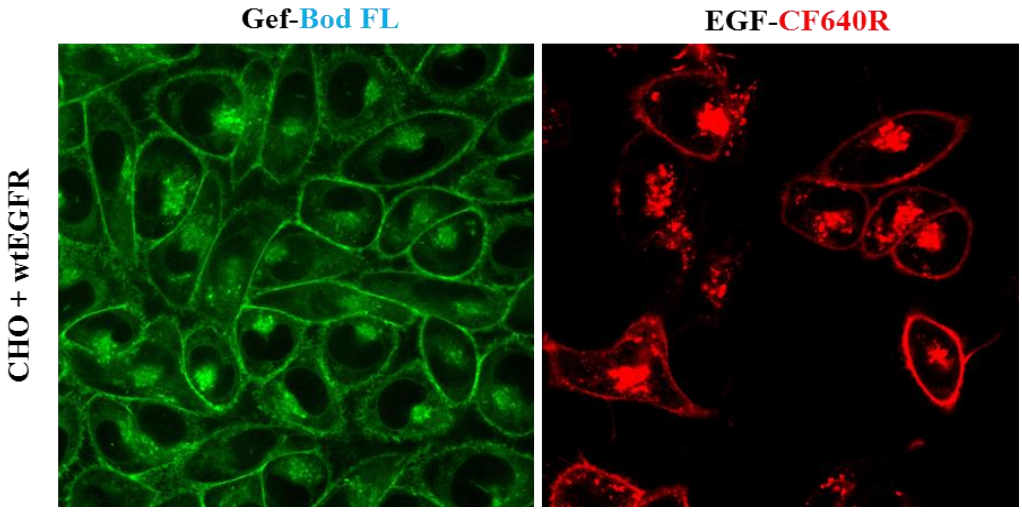


Figure 5-6: Live cell imaging using reversible fTKIs without washing out the labelling media. Live CHO cells expressing wt EGFR were pre-treated with fTKIs for 2 h and then imaged without removing the labelling media. Left image shows the intracellular labelling with a reversible fTKI and the right one the extracellular labelling with EGFR Affibody.

While live cell labelling with fTKIs is very useful for investigating the perturbed receptor, it is not suitable for studying the receptor in its unperturbed state (basal or activated), which was one of the aims of the current project. Therefore, a different labelling method needed to be developed.

In order for the fTKIs to tag their target without perturbing the system cells have to be fixed in advance. Moreover, in order to avoid internalization of the receptors the labelling of the ECD was performed at 4^{o}C. Cells expressing wt EGFR were therefore labelled with either Affibody or EGF and then fixed with 3% PFA and 0.5% GA at the 4^{o}C for 15 min and then at room temperature for another 15 min. They were also briefly permeabilized (only for 1 min) using 0.25% Triton X-100 in order to allow ATP release and hence faster binding of the inhibitors. The results of this labelling method are shown in ***Figure 5-7A***.

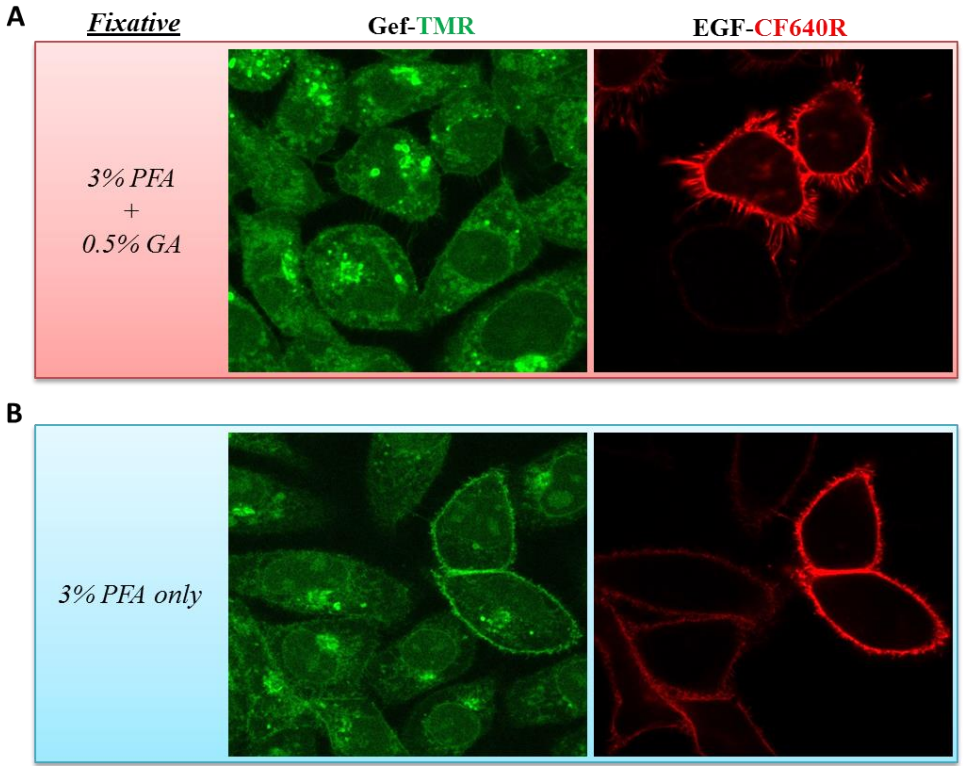


Figure 5-7: Comparison of fTKIs binding in wt EGFR expressing cells pre-fixed with 3% PFA + 0.5% GA (A) or with 3% PFA. CHO cells expressing wt EGFR were labelled with an ECD probe (Affibody or EGF) and then fixed as indicated in (**A**) and (**B**). Cells were then briefly permeabilized and treated with fTKIs before imaging. For both (**A**) and (**B**), left images show the intracellular labelling with Gef-TMR and right ones the extracellular labelling with EGF.

In contrast to live cell labelling, the labelling efficiency of both reversible and irreversible fTKIs in pre-fixed cells was very similar. Another benefit of this method is that impermeable Gef- and Pel-TMR probes can also be used. However I noticed that the signal on the membrane in the fTKIs channel was not comparable to ECD labelling.

I speculated that the presence of GA, which is a powerful crosslinker, in the fixative might crosslink the ATP binding sites. Such a crosslinking could either prevent the release of pre-bound ATP molecules or completely block substrate access. In order to test the former hypothesis, cells were first lightly fixed using 1% PFA for 15 min, then permeabilized (which could facilitate ATP release) and then fixed for another 15 min using 3% PFA+0.5% GA. However, there was no difference in fTKI labelling between this method (data not shown) and the method mentioned earlier (fixation with 3% PFA and 0.5% GA directly). Consequently I decided to remove GA from the fixative to test the latter hypothesis. The labelling method and the concentrations of the probes were exactly the same as described above, except for the composition of the fixative. This

time the fTKI membrane signal was significantly increased and it was also proportional to and colocalized with ECD labelling (**Figure 5-7B**). Furthermore, the background labelling in cells fixed with PFA appears to be lower compared to fixation with PFA+GA (same concentration and incubation time of Gef-TMR were used). Even though is not possible to define the mechanism by which GA prevents substrate binding, these experiments show that GA blocks the ATP sites.

In the process of developing the method for the labelling of pre-fixed cells I noticed that different fluorescent dyes conjugated to the same TKI resulted in fTKIs with different intracellular distribution patterns (***Figure 5-8***). Nevertheless, I do not anticipate issues from these differences as I am specifically interested in the conformational and signalling dynamics of plasma membrane EGFR whose labelling is unaffected.

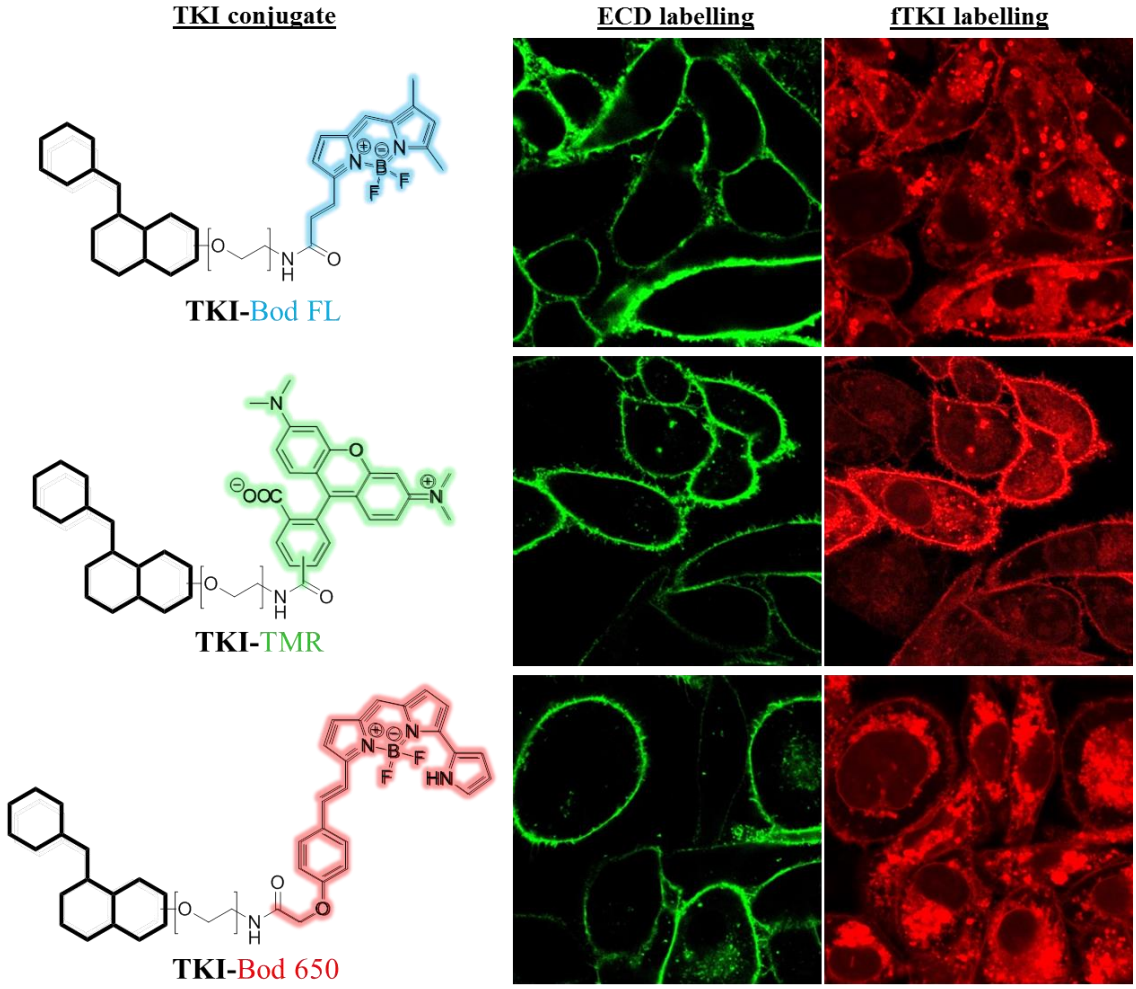


Figure 5-8: Effect of fluorophore on fTKIs' intracellular distribution. Pre-fixed wt EGFR expressing cells prepared as described in ***Figure 5-7B*** were treated with fTKIs and imaged. On the left, the general structures of TKI

conjugates are shown. On the middle and right, examples of extracellular and intracellular labelling, respectively, are shown.

Out of the three different fluorophores, the conjugates with the highest membrane signal-to-background labelling ratio were the TMR labelled compounds (***Figure 5-8*** & ***Figure 5-9A***) whereas the Bod 650 conjugates had the highest background labelling. The fact that the same was true for all TKIs tested suggested that it was not an artefact of a certain conjugate but a general pattern due to dye properties. In fact, it is known that due to their lipophilicity Bodipy dyes have the tendency to accumulate into different subcellular organelles[204] or to form aggregates[205]. Furthermore, Bod 650 small-molecule conjugates have been previously reported to label the endoplasmic reticulum (ER)[187,188] which is what it was observed. Therefore, I decided to carry out the rest of my experiments on pre-fixed cells using the TMR labelled TKIs.

The next step in the evaluation of the *in vitro* binding was to determine the binding specificity of the fTKIs in the CHO cells. The expression of the receptor in the CHO cell lines chosen for this study is induced and controlled by doxycline[206] and hence without it cells do not express any receptor. The same protocol as for pre-fixed cell binding but without including doxycycline in the media was therefore followed. Labelling with ECD ligand this time serves as a marker for endogenous receptor expression, if any. fTKI concentration and incubation times were exactly the same in all experiments. As shown in ***Figure 5-9B***, there was no detectable membrane labelling in either the ECD or the ICD channel demonstrating the specificity of the fTKIs to EGFR only.

I also designed another control experiment to investigate the specificity of the compounds. This time EGFR expression was induced, as in the test experiments, but the cells were treated for 2 h with a saturating concentration of commercial (unlabelled) Neratinib prior to ECD labelling and fixation. Being an irreversible binder, Neratinib permanently blocks all the available ATP-binding sites and hence fTKIs binding should not be possible. As expected, Neratinib treatment completely abolished membrane signal in the fTKIs channel (***Figure 5-9C***), similarly to non-expressing cells.

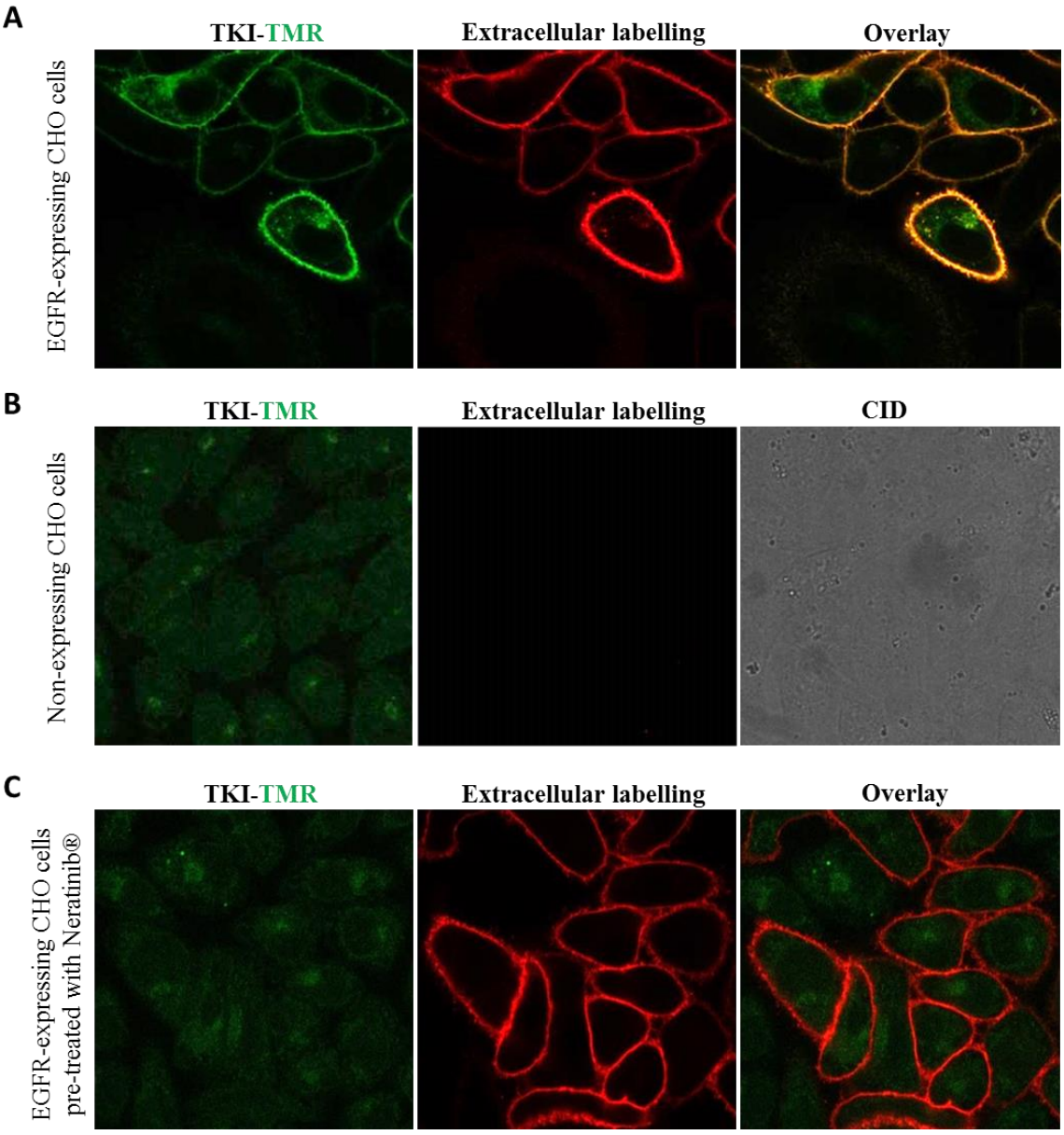


Figure 5-9: *In vitro* binding specificity of TKI-TMR compounds in pre-fixed cells. Comparison of TKI-TMR conjugates treatment in pre-fixed wt EGFR-expressing CHO cells (**A**) with treatment either in non-expressing CHO cells (**B**) or in EGFR-expressing CHO cells blocked by pre-treatment with an excess of irreversible Neratinib® (**C**).

The last part of the evaluation was to assess the *in vitro* binding of the fTKIs against the other ErbB family members. As stable cell lines expressing ErbB2-4 receptors are not currently in use by my group, I transfected wt CHO cells with the desired plasmid using ViaFect (Promega®) and a reverse transfection protocol.

Among the three receptors, ErbB2 transfection resulted in expression levels similar to stably EGFR-expressing CHO cells. For both ErbB3 and ErbB4 the expression levels were generally low despite the efforts to improve them. In the following experiments,

concentrations of all fTKIs were kept the same as in EGFR experiments in order to compare the *in vitro* specificity with the specificity suggested biochemically (K_d). For ECD labelling the following ligands (labelled with fluorescent dyes) were used.

Table 5-4: List of ligands for ECD labelling of ErbB2-ErbB4 receptors

Ligand	Type	Target	Fluorescent dye	Manufacturer
Anti-ErbB2 Affibody ZErbB2:477 (m)	Non-activating	ErbB2	Alexa 488 or CF640R	*Affibody Inc.*
Anti-ErbB3 Affibody	Non-activating	ErbB3		*Gift of Dr J. Lofblom*[207]
Neuregulin-1β (NRG1β)	Activating			*Peprotech*
Anti-ErbB4 Adhiron	Non-activating	ErbB4		*Dr D. Tomlinson (University of Leeds)*

Starting with ErbB2 expressing CHO cells, fluorescently labelled Lapatinib and Neratinib exhibited good labelling efficiency for the receptor, as expected from their low-nanomolar affinities. Membrane signal was also observed in cells expressing ErbB2 after treatment with labelled Gefitinib and Pelitinib, however it was only marginally higher than the background labelling (***Figure 5-10***, blue framework). Among the four TKIs conjugates, only f-Ner was expected to bind ErbB3 and that was consistent with my *in vitro* experiments (***Figure 5-10***, green framework). No detectable signal was observed for any of the other fTKIs for either ErbB3 Affibody or NRG1β labelled receptors. Lastly, ErbB4 labelling was achieved with three out of four fTKIs. Fluorescently labelled Pelitinib and Neratinib showed moderate to good labelling, whereas f-Lap labelling was generally low. No membrane signal was detected from f-Gef treatment (***Figure 5-10***, red framework).

Overall, in the evaluation section I showed that the synthetic strategy for developing fTKIs yielded probes that retained their low nanomolar affinities and selectivity against all ErbB members (both in biochemical assays and *in vitro* binding studies). I also demonstrated that the permeable, irreversible fTKIs can engage and successfully label their target in living cells, whereas there is no limitation in the use of any fTKI with the method I developed for labelling the receptors without perturbation (fTKI labelling after fixation).

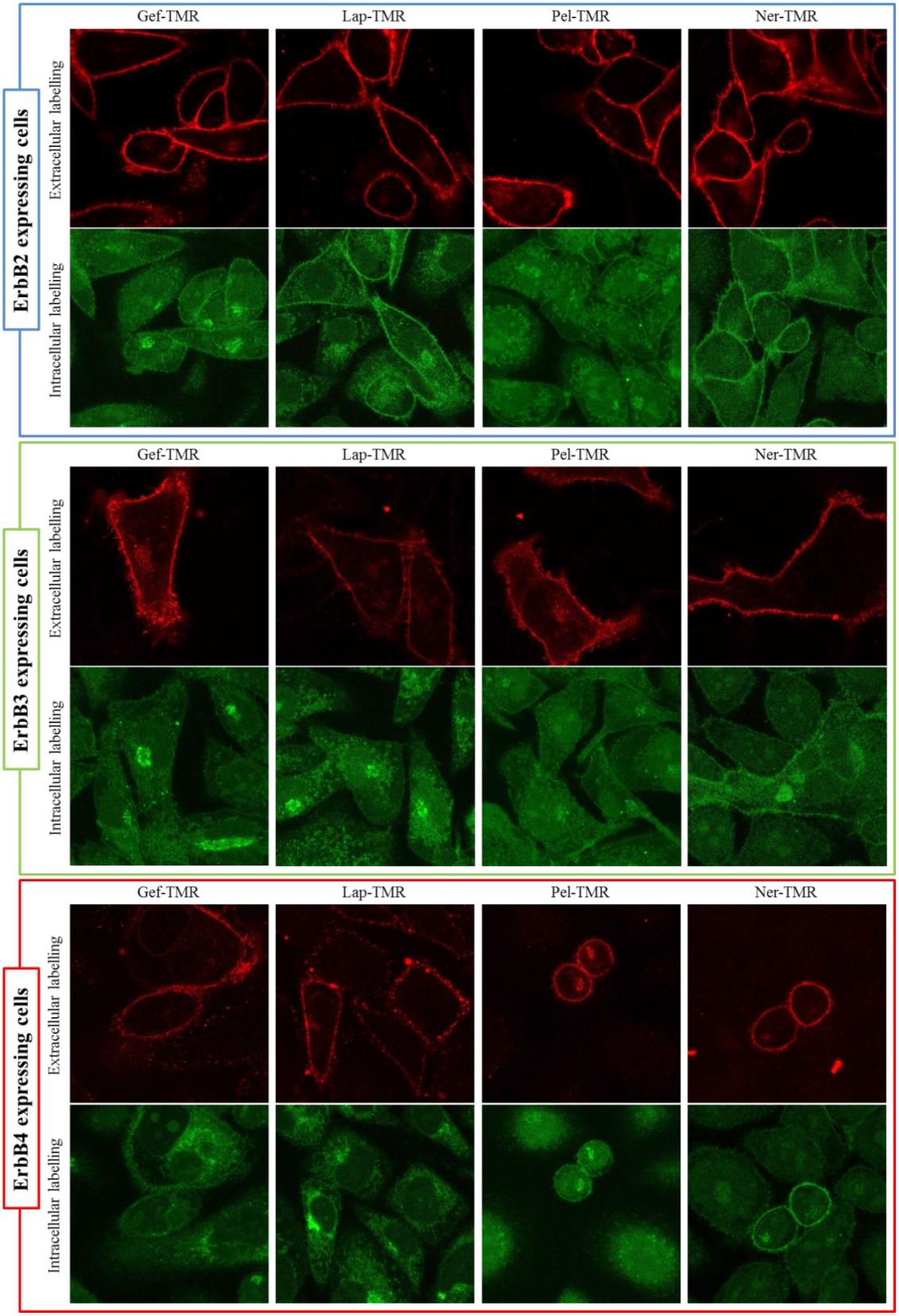


Figure 5-10: fTKIs binding in pre-fixed CHO cells expressing ErbB2-4. fTKIs binding visualization by confocal microscopy in pre-fixed CHO cells (prepared, by analogy, as in ***Figure 5-8B***) transfected with ErbB2 (blue box), ErbB3 (green box) or ErbB4 (red box). ECD probes used for labelling the different receptors are shown in ***Table 5-4***.

5.3 Application of fTKIs

5.3.1 Quantitative Binding Analysis

As discussed earlier (***Section 4.2***), one classification of different EGFR TKIs (and generally for all kinase inhibitors[22]) is based on their type. Type I inhibitors are compounds that bind to the active kinase with a DFG-in conformation whereas type II bind to the inactive kinase (an additional pocket is created within the ATP site) with a DFG-out conformation[208]. It is important to note that while EGFR inhibitors such as Gefitinib, Erlotinib, Pelitinib and Afatinib (among others) are classified as type I, Lapatinib and Neratinib (like the rest of the TKIs with a bulky substituent on the aniline ring) meet only one of the two criteria for type II[174,208] inhibitors despite being mentioned as such in literature[209,210]. These two TKIs bind the inactive kinase but with a DFG-in conformation which are characterized by some as type II-like TKIs[210] or as type I½, a terminology which was introduced by Zuccotto *et al.*[211].

The classification of EGFR TKIs, which is based on crystallographic data[25,46,47,212–214], led to the assumption that these TKIs are selective for the state of the kinase (inactive or active) they bind. While for type II TKIs this preference is generally accepted due to their structural requirement for the additional pocket (red framework, ***Figure 5-11A***) created in the inactive receptor, the selectivity of type I inhibitors to the active remains controversial.

Initially, Yun *et al.*[25] suggested that Gefitinib and AEE788[175] (another type I TKI) show a preference for the active receptor by demonstrating that both compounds have higher affinities against a receptor with an activating mutation (L858R) compared to wt EGFR, which was assumed to be in the inactive conformation in the absence of ligand. Qiu *et al.*[215], and subsequently Wang *et al.*[216], later demonstrated (using an assay developed in the former's lab) the preference of Erlotinib for the active (EGF-treated receptor) conformation of near full-length EGFR alleles (including wt EGFR) compared to the inactive one (Cetuximab-treated receptor).

On the other hand, many biochemical binding studies using TKDs of wt EGFR and activating mutants revealed that both TKI types bind the receptors with similar affinities[176,182,183]. Furthermore, Lu *et al.*[217] showed that the binding affinities of both

Gefitnib and Erlotinib for an EGFR mutant which cannot adopt the active conformation (V924R) due to a mutation on the asymmetric dimer interface are comparable with their affinities for wt EGFR. It is worth noting that these data constitute the only data that can be found in literature regarding the affinity of a compound against an activation-impaired receptor. Based on this study, Park *et al.*[218] performed molecular docking of Erlotinib against the V942R mutant and showed that the Glide score (a measure of binding affinity for *in silico* studies) is similar to that obtained by docking Erlotinib against the active receptor. Most importantly though, they resolved the crystal structure of Erlotinib in complex with V924R which is the only structure where a type I TKI was seen bound on an inactive kinase conformation (***Figure 5-11C***).

Despite these findings, the question regarding the preference of type I TKIs is still not clear[218]. The necessity however for understanding it derives from the clinical need to pre-select drugs with the highest possible impact against different EGFR alleles and thereby overcoming inhibitor resistance.

For this purpose, I sought to exploit the fTKIs to unravel the mechanisms of action of TKIs. The main advantage over the studies referenced above (biochemical and crystallographic) is that the post-fixation fTKI labelling method allows to directly monitor the binding of the inhibitors to the receptor in their cell environment.

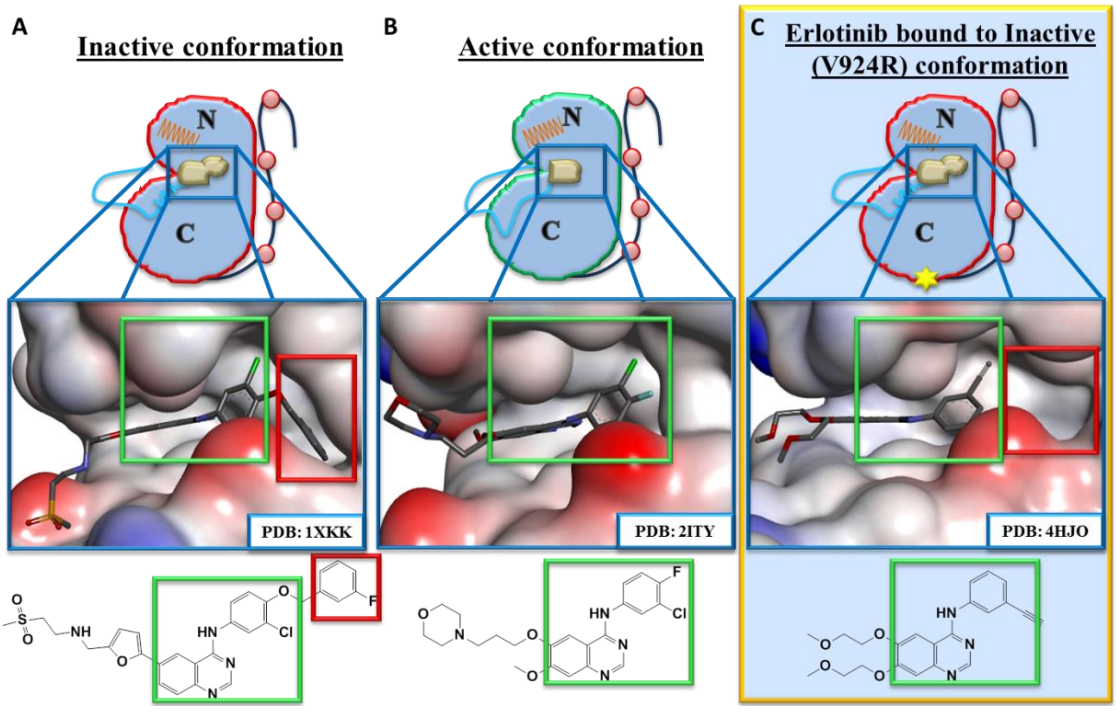


Figure 5-11: Conformation of the ATP cleft in the inactive and in the active kinase domain and its occupancy by TKIs in each state. As crystal structures have revealed, type II TKIs (here Lapatinib in PDB 1XKK, **A**) require

an additional pocket, compared to the active conformation (PDB 2ITY, **B**), to be formed deep in the cleft in order to be accommodated in the ATP-site (inactive conformation). In the wt receptor, both type I and II TKIs fully occupy the ATP cleft. On the other hand, binding of type I TKIs in a receptor which is unable to be activated (e.g. V924R) leaves the addition pocket present in the inactive conformation unoccupied (**C**).

In order to determine whether type I TKIs have any such preference or not I employed the pre-fixed cells labelling method (***Figure 5-9***) combined with quantitative analysis. It is known that type I TKIs promote, irrespectively of their preference to any state, the active conformation[65,217,219] (similarly, type II promote the inactive one) and therefore utilization of this method would allow me to monitor fTKIs binding in either the basal (inactive) or the EGF-treated (active) receptor without allowing the inhibitors to induce any conformational change that would bias my results, since cells would be pre-fixed.

The experiment I designed was as follows: wt EGFR-expressing cells were left untreated (basal-inactive) or were activated with a range of EGF concentration for 1 h in at 4^{o}C before fixation. Cells were then treated with fTKIs and data was acquired without removing the fTKI-containing media for 25 min. The receptor expression levels per cell were measured using a nanobody (EgB4[220], gift of Prof van Bergen en Henegouwen) which does not activate the receptor and whose binding is unaffected by EGF binding as it binds to a different epitope (ECD sub-domain I)[220]. Labelling with EgB4-CF640R was performed simultaneously with (or without) EGF.

Quantification of the binding was performed using an in-house developed plugin (Dr Chris Tynan) for ImageJ. EgB4-AF488 pixel intensity (EGFR expression) and the mean intensity of the corresponding pixels in the fTKI channel (inhibitor binding) for EGF concentrations in the range 0-20 nM were plotted (***Figure 5-12A***). Simple linear regression was used to quantify statistical relationships between EGFR expression and inhibitor binding. The gradient of the best fit lines for each EGF concentration is plotted in ***Figure 5-12B***.

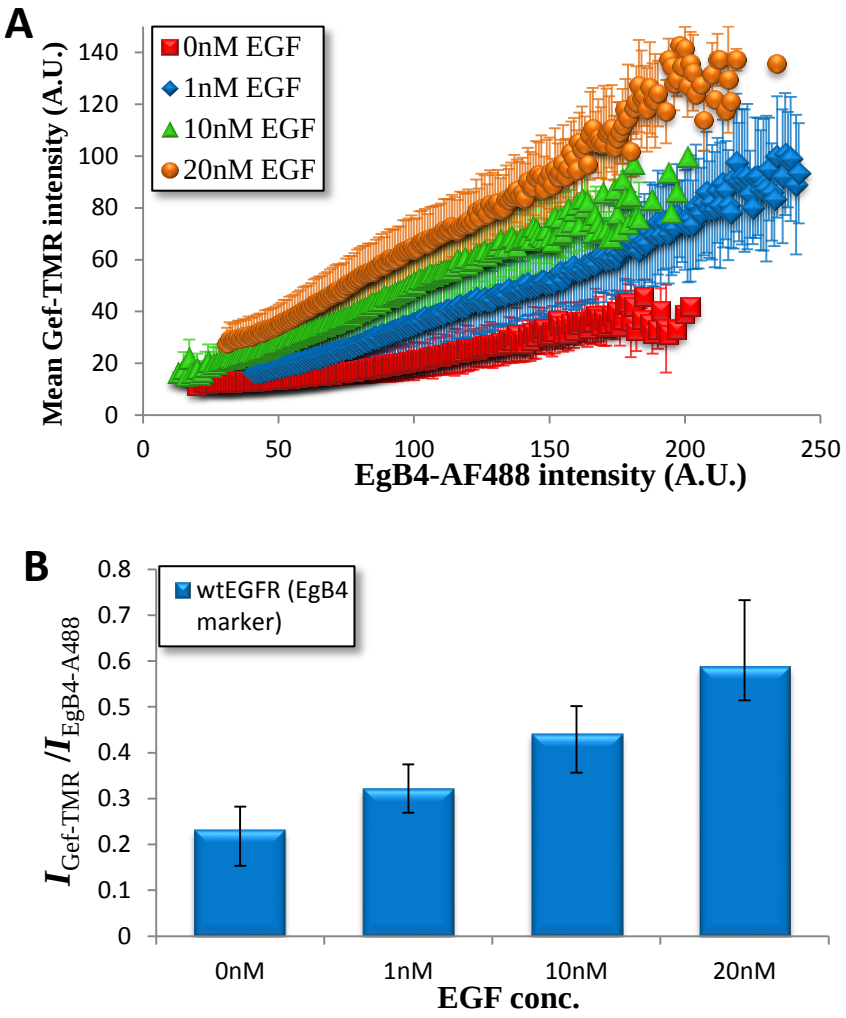


Figure 5-12: Relationship between EGFR expression and fTKI binding exposed to invreasing EGF concentration (A) and the gradient of the best fit lines for each EGF concentration (B). CHO cells stably expressing wt EGFR were treated with a mixture of 50 nM EgB4-AF488 (labelling probe) and with or without EGF, fixed with 3% PFA and then briefly permeabilized. Cells were then treated with 100 nM Gef-TMR and data were collected without removing the labelling media. All experiments were performed in triplicate. (**A**) EGFR expression (EgB4-AF488) and fTKI binding (Gef-TMR) in different EGF concentrations (0-20 nM) were quantified by plotting EgB4-AF488 pixel intensity with the mean intensity of the corresponding pixels in the fTKI channel for all four conditions. (**B**) The gradient of the best fit lines from (**A**) were plotted in order to quantify the statistical relationship between EGFR expression and fTKI binding for each EGF concentration.

The data for the Gef-TMR experiments show that Gefitinib binding in wt EGFR-expressing cells increased with increasing EGF concentration. It is known that the activation of the wt EGFR receptor is EGF-dependent and also that receptor activation occurs through the formation of an asymmetric dimer in which the individual TKDs are in the active conformation. Therefore, the increased fTKI affinity observed in these experiments is conformational-dependent. The higher the concentration of EGF the

more “active” kinases are generated which subsequently leads to higher Gefitinib binding.

In order to validate my interpretation regarding the increasing number of “active” kinases with increasing EGF concentration, the phosphorylation levels of the receptor when exposed in increasing EGF concentrations were measured by Western Blot (WB). The more “active” TKDs in the receptor population, the higher the phosphorylation of the receptor should be. The protocol use for the WB was identical to the fTKI method mentioned earlier up to the fixation step, at which cells were lysed instead (followed by gel electrophoresis and blot development). WB analysis showed that the phosphorylation of wt EGFR is increased with increasing EGF concentrations (***Figure 5-13***), supporting the hypothesis that the increased Gef-TMR binding is due to increasing number of active kinases.

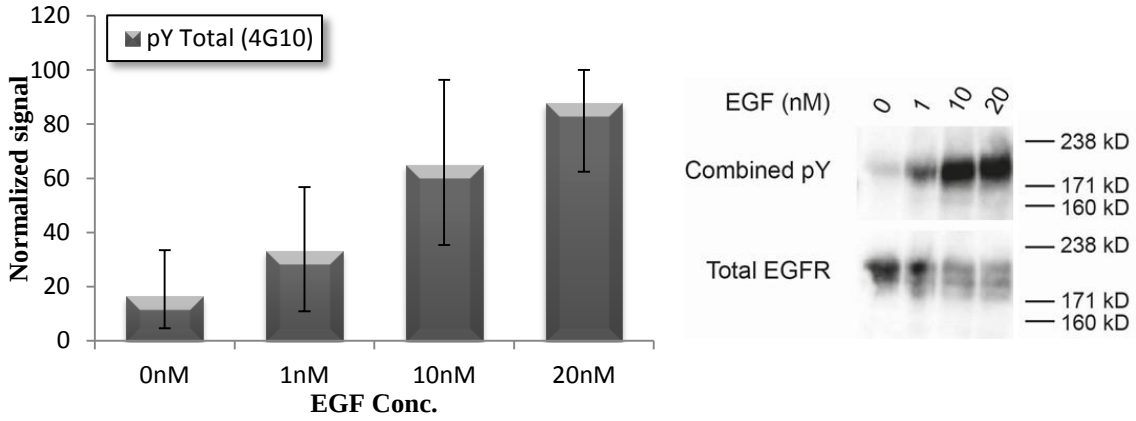


Figure 5-13: Western blot measurement of total tyrosine autophosphorylation of wt EGFR in CHP cells exposed to increasing EGF concentration. Western blot results using the pan-phosphotyrosine antibody 4G10 and an antibody against the EGFR N-terminal tail from CHO cells expressing wt EGFR. EGFR phosphorylation was measured in cells treated with increasing EGF concentration (0 (untreated), 1, 10 or 20 nM) as described in ***Section 3.5.3***. All experiments were performed in triplicate. The phosphorylation signals are normalized to total EGFR signal and the mean and standard deviation of three replicates has been plotted

Furthermore, the WB analysis showed that the receptor is phosphorylated even in the absence of EGF and that this phosphorylation is ~5-fold lower than signal at 20 nM of EGF. The fact that the binding of Gef-TMR in the basal state is only ~2.5-fold lower than the binding at 20 nM of EGF indicates that only part of the basal binding can be attributed to binding to inactive receptors (monomers or symmetric dimers) as the rest is due to binding to the active/phosphorylated population. Therefore, Gefitinib affinity for

the inactive kinase is lower compared to the affinity for the active kinases, clearly suggesting a preference towards the active conformation.

Having explored the preference of type I reversible TKIs, I sought to use the type I irreversible Pelitinib in order to further understand their mechanism of action. The results are shown in ***Figure 5-14***. In contrast to Gefitinib, Pelitinib affinity appears to be independent of receptor activation in all the EGF concentrations tested. The covalent binding can therefore alter the preference of type I reversible TKIs. This offers an insight into the kinetics of TKIs preference.

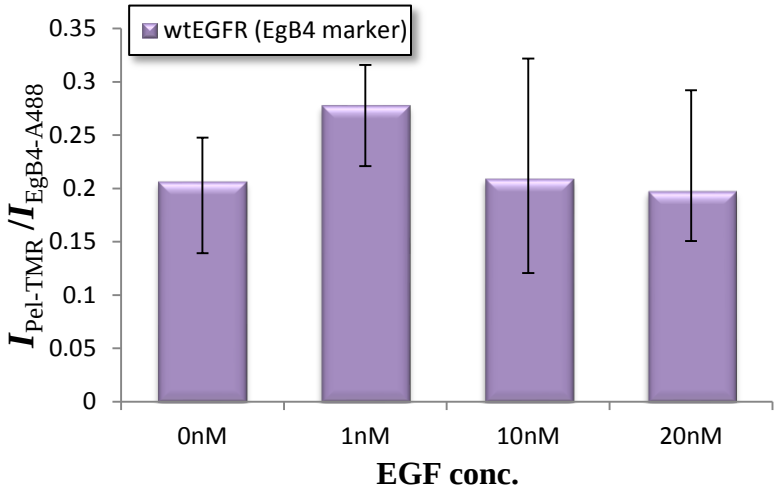


Figure 5-14: Pel-TMR binding in fixed CHO cells expressing wt EGFR exposed to increasing EGF concentration. Cell preparation and data analysis were performed as described in ***Figure 5-12***. All experiments were performed in triplicate.

In the context of experiments where factors such as permeability rate, ATP competition and ATP-pocket desolvation are eliminated, the binding kinetic models could be simplified as illustrated in ***Figure 5-15A***. The affinity of a reversible TKIs depends on the ratio of the dissociation (K_{off}) over the association (K_{on}) rate of the inhibitor towards the target ($K_d=K_{off}/K_{on}$, green rounded rectangle, ***Figure 5-15A***). On the other hand, for irreversible TKIs bearing a Michael acceptor the K_d only depends on the K_{on} as after binding and covalent adduct formation they cannot be dissociated (red rounded rectangle, ***Figure 5-15A***).

Even though Pelitinib has not been crystallized with EGFR, the structural similarity of its anilinoquinazoline core (which is located within the pocket) to that of Gefitnib (blue bold structures, right and left repsectively, ***Figure 5-15B***) suggest that their

binding mode should be very similar[221]. The only difference between them lies in the cyano-group of Pelitinib (red bold structure, ***Figure 5-15B***) which has been proposed to form a hydrogen bond with a nearby threonine residue[185].

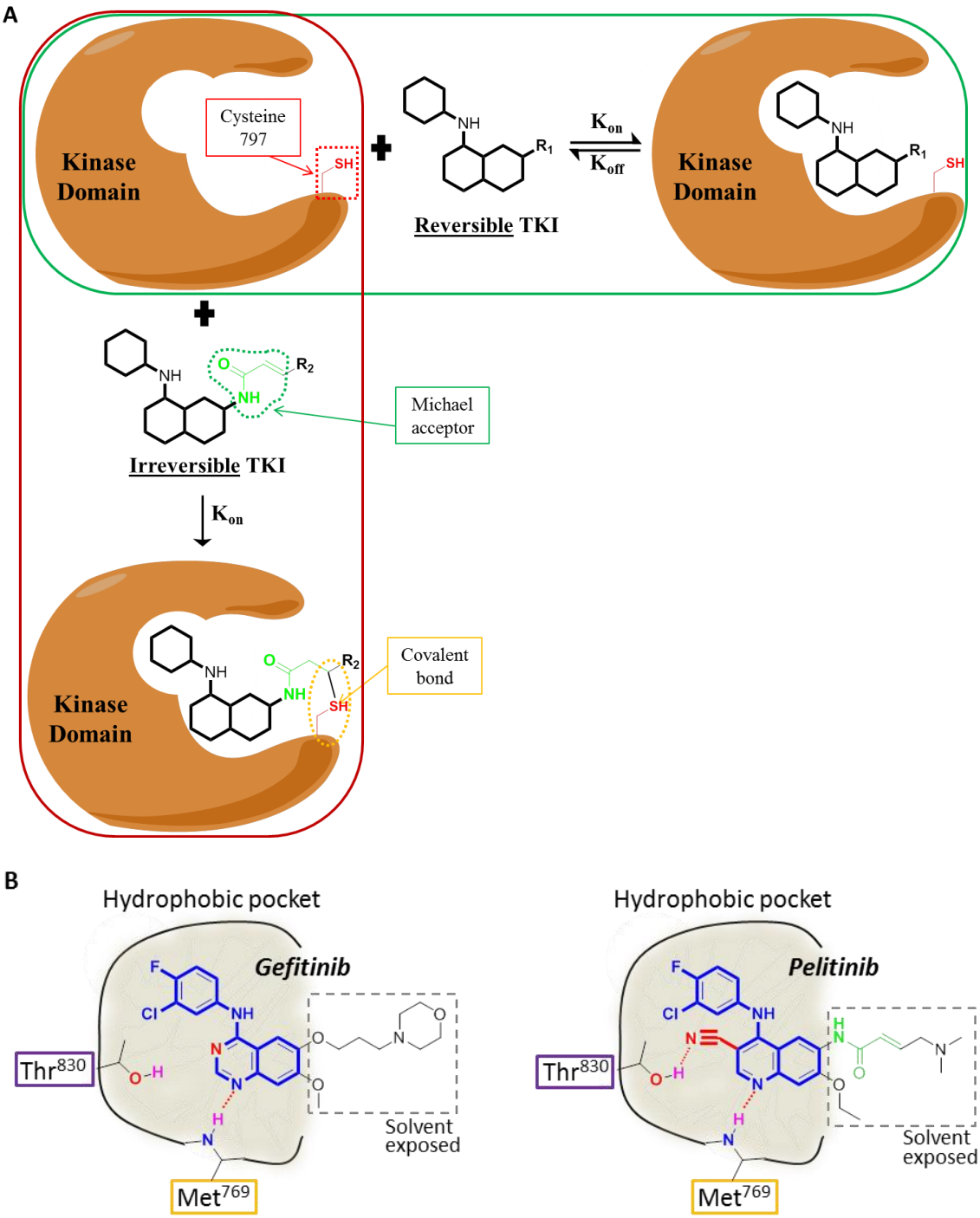


Figure 5-15: Binding kinetics of reversible and irreversible EGFR TKIs (A) and structural and binding mode similarity of Gefitinib with Pelitinib (B). (**A**) Binding affinity (K_d) of reversible TKIs to the TKD depends on both the association (K_{on}) and dissociation (K_{off}) rate (green, horizontal). On the other hand, TKIs bearing a Michael acceptor covalently bind to the TKD and therefore their binding affinity only depends on the K_{on} (red, vertical). (**B**) The structural similarity of Gefitinib (left) and Pelitinib (right) cores which are placed within the ATP pocket (blue bold) suggests that both bind in a similar way.

Taking into account the kinetics and the binding similarity of Geftinib and Pelitinib, the data suggest that the reduced affinity of type I reversible TKIs against the inactive EGFR is due to their faster dissociation (K_{off}) from the pocket and not due to their inability to associate (K_{on}), otherwise Pelitinib would also have lower affinity in the basal state.

In order to further support my interpretation regarding the dependency of Gefitinib binding on the conformational changes occuring with increasing EGF concentrations, I employed the type II fTKIs Lapatinib (reversible) and Neratinib (irreversible). As discussed earlier, type II TKIs are not expected to bind to the active conformation due to their structural restrictions (they cannot be accommodated in the "small" active site) and hence the increase in the number of active kinases with increased EGF in my experiments should result in reduced bindings compared to untreated cells.

The results from the experiments with both type II fTKIs, performed using the same method as earlier, are shown below (***Figure 5-16***). In both cases, binding was decreased upon treatment with EGF and the decrease became more pronounced when EGF concentration was increased. This shows that the affinity of type II TKIs, irrespectively of whether they are reversible (Lapatinib, ***Figure 5-16A***) or irreversible (Neatinib, ***Figure 5-16B***) binders, is EGF-dependent and that this dependency is inversely proportional to receptor activation/phosphorylation, as opposed to type I. Furthermore, the fact that both type II inhibitors behaved in the same way in response to EGF suggests that the parameter responsible for the reduced binding is the K_{on} due to steric hindrance from the smaller ATP-site in activated receptors

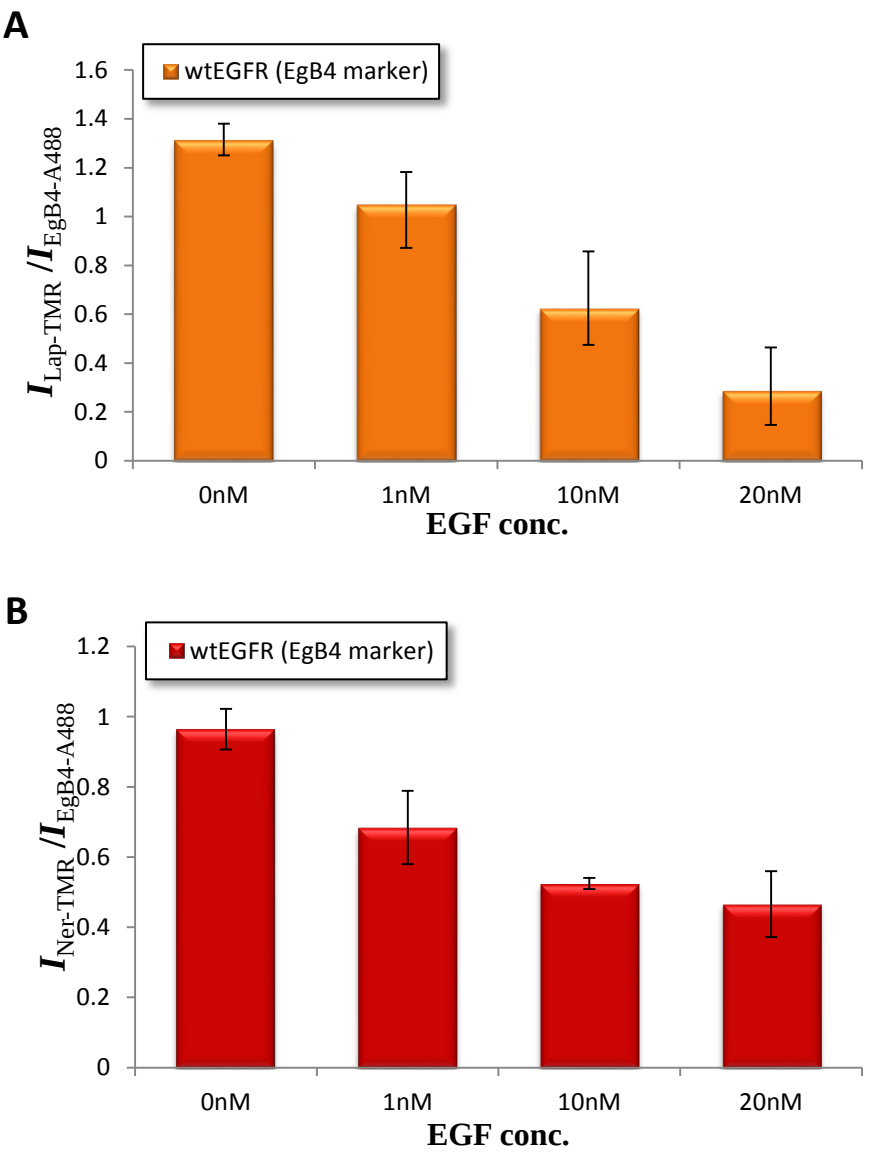


Figure 5-16: Type II fTKIs binding in fixed CHO cells expressing wt EGFR exposed to increasing EGF concentration. Determination of f-Lap (**A**) and f-Ner (**B**) binding in wt EGFR expressing CHO cells prepared. Cell preparation and data analysis was performed as in ***Figure 5-12***. All experiments were performed in triplicate.

In conclusion, using these novel tools and method I was able to show not only that type I reversible TKIs do have a conformational preference towards the active kinase (higher binding in EGF-treated cells), but also that this preference is due to faster release of the inhibitors from the inactive kinase (differences in the binding between reversible and irreversible type I fTKIs), despite their ability to associate with it. Even though this association is in agreement with the findings from Park *et al.*[218] (they demonstrated the ability of Erlotinib to bind an inactive receptor by crystallization), the type I reversible TKIs' preference shown herein is in contrast with their conclusion

(type I TKIs bind both conformations with same affinity)[218]. Even though a crystal structure of a bound inhibitor is informative for its binding mode and its interactions, it does not allow getting an insight into the affinity of a compound. For instance, Gajiwala et al.[222] resolved the structure of Gefitinib with a TKI resistant double EGFR mutant (L858R/T790M), for which Gefitinib is known to have very low affinity[176] and also not to be able to inhibit its phosphorylation[223]. If anything, the crystal structure of Erlotinib bound to the inactive kinase might suggest a lower affinity compared to its binding to the active. An explanation of this view might reside in the fragment-based approach used in drug discovery whose strategy is, briefly, to create high-affinity compounds by filling in the empty gaps in a binding pocket[224]. The fact that the binding pocket of the inactive kinase is not fully occupied by Erlotinib (***Figure 5-11C***), as in the case of Lapatinib (***Figure 5-11A***), is reminiscent of the concept of this strategy and my data is in agreement with this hypothesis.

Another important outcome from these experiments is that fTKIs can be used as reporters of the conformation/activation of the receptor. This would allow me to determine the structural-functional relationship of EGFR in cells when combined with existing imaging techniques.

5.3.2 Intradimer FRET between fTKIs

As discussed earlier (***Section 1.2***), activation of ErbB receptors is inextricably linked to dimerization, irrespective of whether oligomerization/clustering of EGFR is also involved, as asymmetric dimers are the smallest known signalling units required for phosphorylation to occur. For that reason, receptor dimerization was initially thought to be an event that only occurs after activation[225]. However, early studies suggested the presence of pre-dimers even in the absence of ligand[101,226]. Later independent studies also proposed that ErbB family members pre-dimerized[66–68,225]. Furthermore, formation of pre-dimers in the basal receptor state has also been suggested to be associated with EGFR activation[66,68,227]. While asymmetric dimerization and phosphorylation after EGF stimulation is well documented[49,54,63,65,228,229], the presence of such pre-dimers is still controversial[62,65] (***Section 1.2.2***). For these reasons, monitoring dimerization events and the extent to which they occur is essential for understanding the mechanisms of EGFR activation. Furthermore such a method could potentially allow us to assess the

prognostic significance and predictive value of HER receptor dimerization in various cancers[230].

The most frequently employed methods for identifying protein interactions are FRET-based or biochemical methods involving chemical cross-linking[231]. The main disadvantages of using the latter approach are the requirement for cell disruption and the inability to preserve all dimeric structures[232,233]. On the other hand FRET experiments can be performed in both fixed and live cells and allow the distance between FRET pairs to be estimated under carefully controlled conditions.

FRET-based methods used for studying the dimerization of ErbB family members (either before or after activation by ligand) include, steady state FRET (using both confocal[234] and TIRF microscopy[73,101]), FLIM-FRET[230,235] and homo-FRET[236–238]. Labelling of the receptors in these studies has been performed using labelled molecules (ligands, antibodies, etc) targeted to the ECD, labelled antibodies for the ICD, or fusion with FPs. However, these methods of labelling have drawbacks for studying interactions between EGFR receptors. For instance, the large size of antibodies runs the risk of producing a FRET signal due to crowding rather than direct interactions. The use of labelled ligands such as EGF is not suitable for FRET experiments as the distance between ligand binding sites in an ECD dimer exceeds 10nm, a distance where FRET is either no longer possible or practically undetectable between available FRET pairs[68]. FPs are also large and are usually fused to the ~200 amino acid long and flexible C-terminal tail which might perturb the functioning of the receptor.

Having labelled all the TKIs with more than one colour, I sought to measure FRET between two fTKIs within dimeric receptors. The advantage of this labelling method over the other methods for monitoring EGFR interactions is that the small fTKIs bind to the centre of the TKD and that the distances between fluorophores in either the symmetric or the asymmetric dimers are within FRET range as measured by molecular simulations (Dr Martyn Winn, ***Figure 5-17***). The importance of assessing EGFR dimerization by monitoring TKD-TKD interactions is further highlighted by the fact that EGFR TKDs haves been suggested to be necessary for receptor pre-dimerization[225,239]. For these experiments I used Bod FL and Bod 650 labelled TKIs in order to avoid crosstalk and bleed-through. A FLIM-FRET setup was preferred over intensity-based FRET methods for reasons discussed in ***Section 1.3.1***.

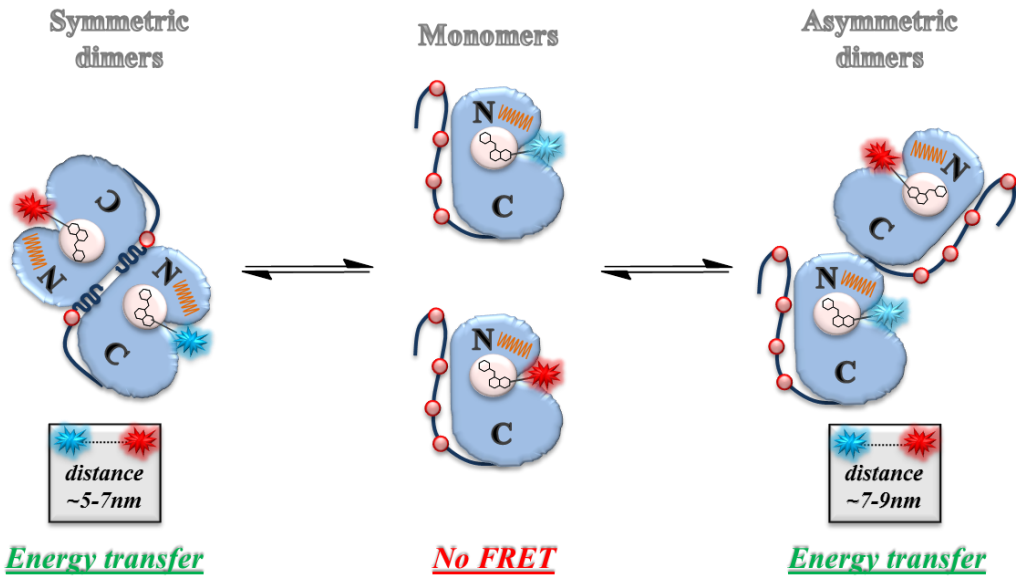


Figure 5-17: Illustration of possible structures and fTKIs distances for FRET experiments.

As discussed in ***Section 1.2.1***, the three possible conformations of EGFR TKDs are: monomers, symmetric dimers or asymmetric dimers. As shown in ***Figure 5-13***, basal wt EGFR is mostly populated by inactive kinases which can either exist as monomers or symmetric dimers. Therefore, type II fTKIs that recognise the inactive conformation or Pelitinib which binds both inactive and active conformation (***Section 5.3.1***) would in principle be able to detect most of the dimeric structures, if any. However, I selected the two irreversible fTKIs for the following experiments as the fact that they cannot be dissociated from the ATP site might increase the chances for detecting FRET.

I started the experiments using Neratinib conjugates in pre-fixed EGFR-expressing cells, as described earlier (***Section 5.3.1***). After allowing 20 min for Ner-Bod FL to bind the receptor, the lifetime (***Figure 5-18***) of donor only samples labelled with four different concentrations of fTKI (10-80nM) was measured. The mean lifetime of the membrane signal in all conditions tested was 5.05 ns. Even though this value is different from the lifetime of the "free" dye in methanol[240](possibly due to the different solvent and the binding of the fTKI conjugate[241–243]), it is consistent with the reported lifetime of Bod FL conjugated to protein (5.0 ± 0.2 ns)[244]. Acceptor (Ner-Bod 650) only samples were also prepared in order to ensure that direct excitation of Bod 650 with the 488 nm laser does not occur (not shown).

As shown earlier (***Figure 5-8***), Bod 650 conjugates tend to accumulate in the ER, despite their ability to selectively bind the receptor. This tendency effectively reduces

the concentration available locally to receptor binding. Moreover, Bod 650 conjugates exhibited a ~2-fold decrease in their binding affinity compared to Bod FL labelled TKIs (***Table 5-1***). As a result, the concentration required to obtain detectable membrane labelling was much higher than for Bod FL conjugates. A high acceptor concentration is also required to reduce the occurrence of donor-donor pairings, which would lower the measured FRET efficiency.

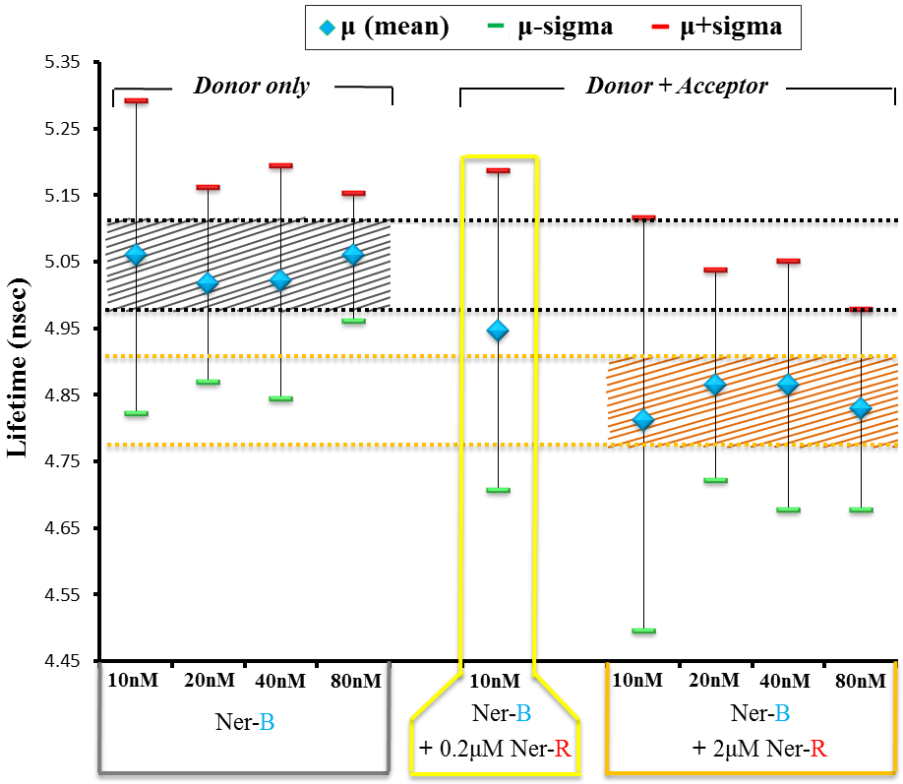


Figure 5-18: Donor fluorescent lifetimes in samples with donor (Ner-Bod FL) only and in samples with increasing concentration of acceptor (Ner-Bod 650). CHO cells expressing wt EGFR were fixed, after 2 h starvation, and briefly permeabilized as described in ***Section 5.3.1***. Cells were then labelled for 20 mins with Ner-Bod FL in indicated concentration. In "Donor only" samples, donor lifetimes (μ) were recorded after removing the labelling media. In "Donor + Acceptor" samples, the donor labelling media was replaced with acceptor (either 200 or 2 μM) labelling media, left for 20 mins and then donor lifetimes were recorded without removing the acceptor labelling media. All experiments were performed in triplicate. Analysis was performed as described in ***Section 3.6.2.1***.

Receptor labelling in the "Donor + Acceptor" (D+A) experiments could be performed by treating the cells with both donor and acceptor simultaneously, the donor first following by acceptor labelling or the acceptor first. However, it was found that by treating cells with the donor first (20 min incubation) following by replacement with the acceptor media (which remained in the dish during acquisition) the equilibrium between

donor and acceptor binding that produced the largest observed donor lifetime reduction was reached faster compared to the other methods. In both the simultaneous labelling and the sequential labelling with acceptor first, the acceptor membrane labelling was not as efficient. The reason for that was probably the slower binding of Bod 650 conjugates (compared to Bod FL ones) due lower effective concentration and competition with Bod FL conjugates, factors which cannot be easily pre-determined.

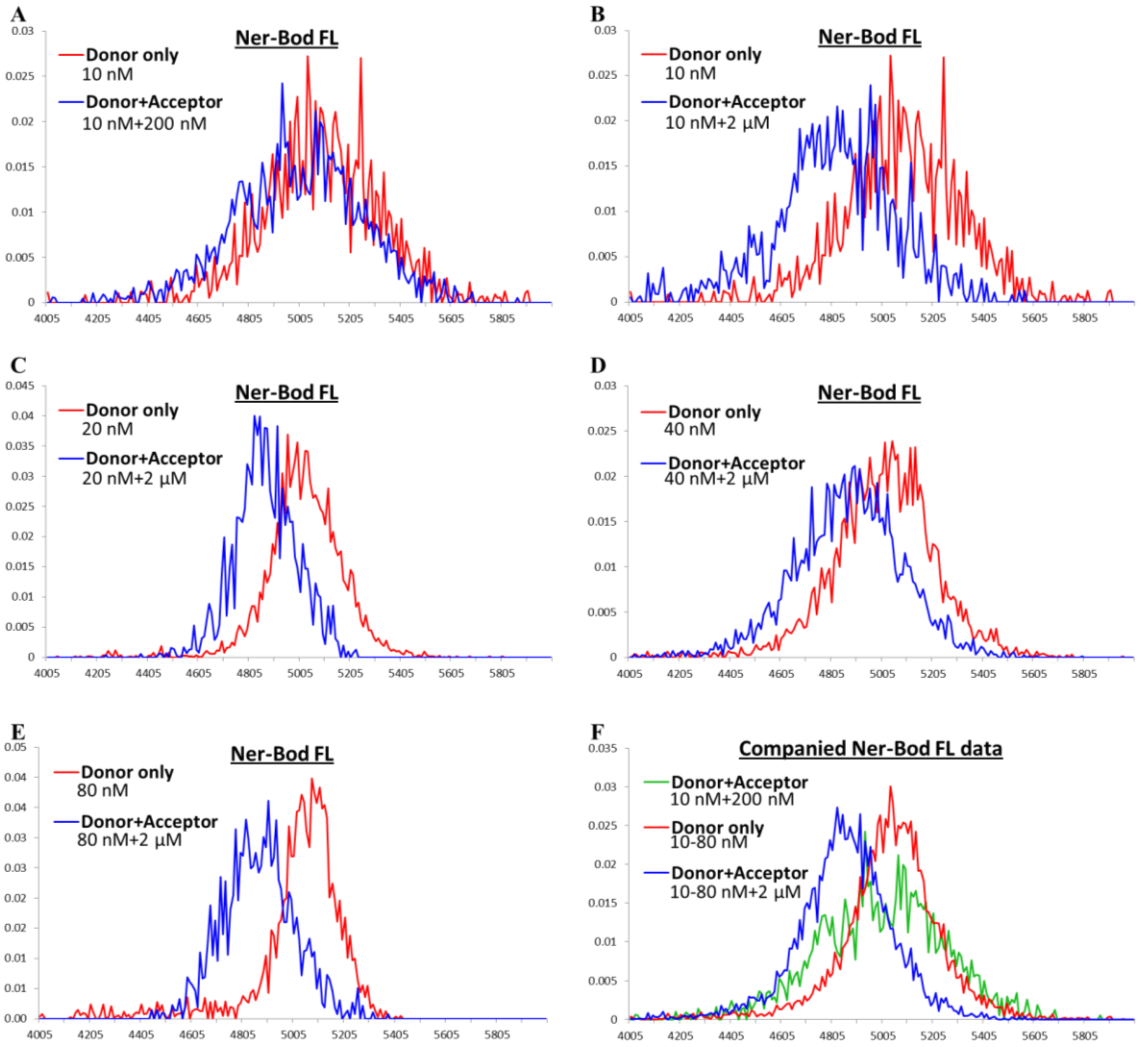


Figure 5-19: Comparison of the probability histograms of the mean lifetimes of "Donor only" with "Donor+Acceptor" samples for each FRET pair concentration. Histograms derived from the lifetime distribution of each condition shown in ***Figure 5-18***. In plots in A-E, the histograms of donor only (red) and donor+acceptor (blue) sample for a given donor concentration are compared. 10 nM of donor with 200 nM of acceptor or 2 µM are shown in (**A**) and (**B**) respectively, 20 nM of donor in (**C**), 40 nM of donor in (**D**) and 80 nM of donor in (**E**). Histograms of the average of lifetimes of all donor concentrations without (red) or with acceptor (green for 200 nM and blue for 2 µM of acceptor respectively) are shown in (**F**).

In the first D+A experiment, 200 nM of Ner-Bod 650 and the lowest donor concentration (10 nM) were used. As expected by the relatively low membrane labelling (intensity images) of the acceptor (bottom middle, ***Figure 5-20***), the mean lifetime across all cell was measured to be only 90 ps lower the donor-only samples (yellow box, ***Figure 5-18***). Comparison of its probability histogram with those of the 10 nM of donor only samples is also shown in ***Figure 5-19A***. Higher donor concentration also tested but completely abolished acceptor binding. Having seen that the acceptor concentration used was not sufficient for membrane labelling I decided to increase it by a factor of 10 (2 μM). This time acceptor membrane binding was significantly increased as shown in ***Figure 5-20*** (bottom right). This increase was also accompanied by a modest reduction, ~200 ps, in the donor lifetime (mean lifetime of ~4.8 ns, ***Figure 5-18***). The same acceptor concentration was also tested with the other three donor concentrations (20, 40 and 80 nM) and showed that the observed lifetime reduction was retained throughout all the different donor concentrations (***Figure 5-18***). This time a clear shift of the lifetime distribution towards shorter lifetimes was observed for all conditions as shown in ***Figure 5-19B-E*** (also in the plot with the combined data, ***Figure 5-19F***).

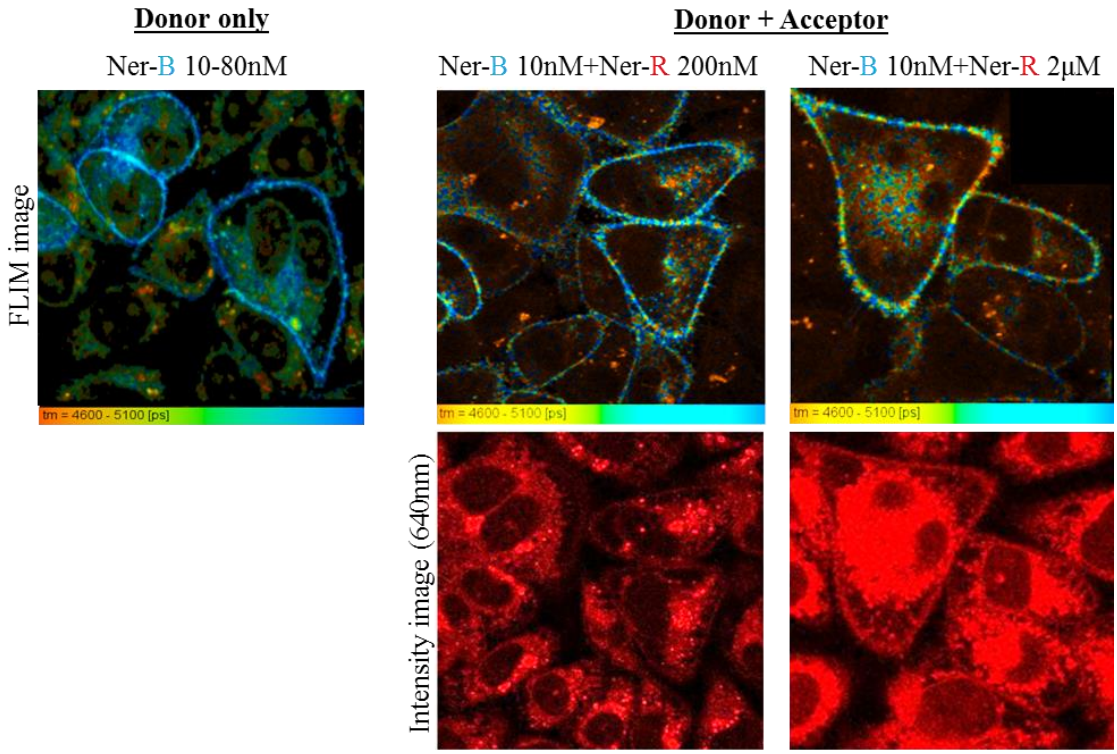


Figure 5-20: Representative lifetime and intensity images of wt EGFR-expressing cells labelled with either Ner-Bod FL only or Ner-Bod FL + Ner-Bod 650. The membrane lifetime of samples labelled with donor only (top

left) is uniformly coloured blue (colour-coded between 4.6-5.1 ns) whereas upon addition of the acceptor (middle and right) the membrane colour becomes variable (same colour-coding as before).

It is worth noting that when higher donor concentrations (160-400nM) were tested, the recorded lifetime of the donor only samples was lower compared to samples with donor concentrations 10-80 nM (***Figure 5-21A***). Since this reduction was not observed with lower concentrations it could be attributed to self-quenching of bound donor molecules due to high concentration[245]. Addition of 2 μM of acceptor in the high donor concentration samples did not result in binding in the acceptor channel (right bottom image, ***Figure 5-21B***) or any further change in the donor lifetime (***Figure 5-21A***). This suggests acceptor associated drops in donor lifetime are not caused by excess unbound inhibitor.

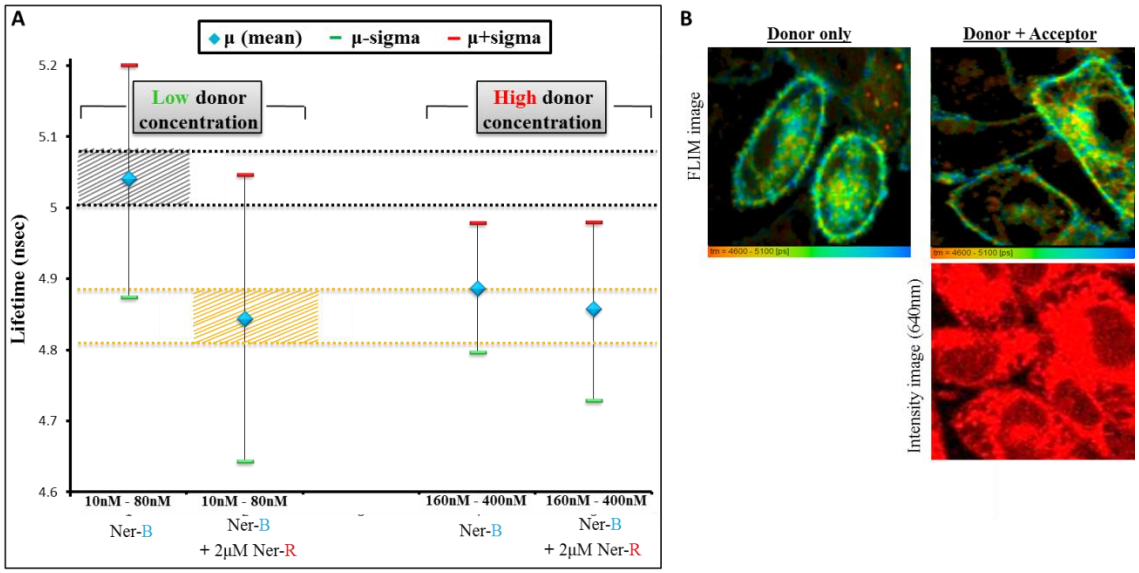


Figure 5-21: Effect of high donor concentration in its lifetime. (**A**) Donor lifetimes of donor, in the absence and in the presence of acceptor, in samples with low (10-80 nM, obtained from ***Figure 5-18***) and high (160-400 nM) Ner-Bod FL concentrations. Samples for the high donor experiments were prepared as described for the low donor concentration (***Figure 5-18***). All experiments were performed in triplicate. (**B**) Representative lifetime and intensity images of wt EGFR-expressing cells labelled with high Ner-Bod FL concentration without or with Ner-Bod 650.

Further evidence of the correlation between acceptor binding and the reduction in donor lifetime is shown in ***Figure 5-22***: when the membrane intensity in acceptor channel was increased (by allowing more acceptor molecules to bind the receptor or by replacing donor molecules from the binding sites) the donor lifetime was further decreased. This suggests that this reduction is indeed due to FRET between donors and

acceptors. Because increasing the donor concentration up to 80 nM did not lead to a further decrease in donor lifetime, it is unlikely that the acceptor induced reduction in donor lifetime is due to bystander FRET.

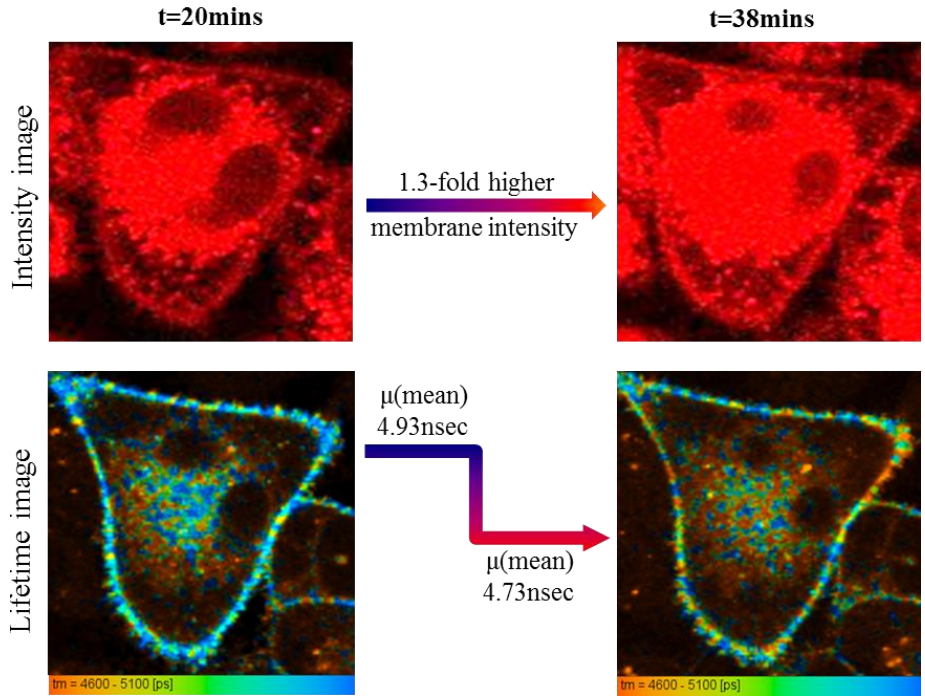


Figure 5-22: Dependency of FRET to acceptor binding. After replacement of the donor with the acceptor (2 μM) media, intensity and lifetime images of the same area were recorded at t=20 min and 18 min afterwards. The increase in the acceptor binding was accompanied by a decrease in donor lifetime, indicating the dependency of FRET on the increased number of acceptor molecules bound to receptors.

It was also observed that the donor lifetime distribution on the cell membrane of D+A samples was not uniform. A close examination revealed that certain areas on some cells exhibited lower lifetime compared to the mean lifetime of the whole cell (***Figure 5-23***). Therefore, FRET events with larger lifetime reduction may occur but are averaged out.

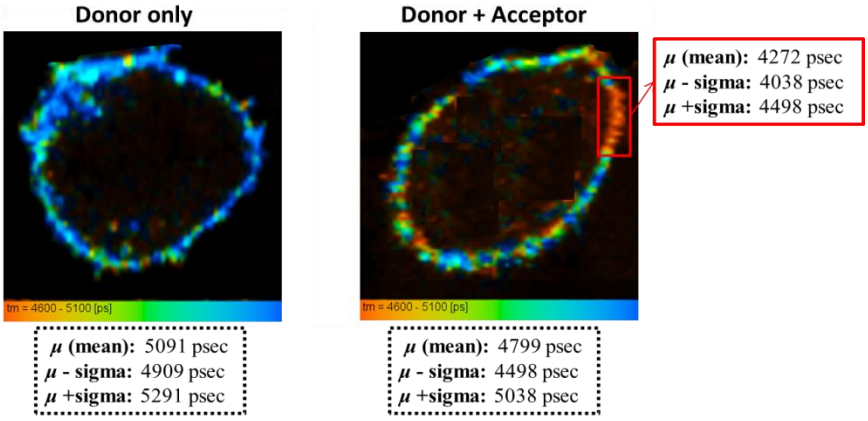


Figure 5-23: Fluctuations of donor lifetime on cell membranes of samples treated with donor only (left) and donor+acceptor (right).

Next I attempted to measure FRET using labelled Pelitinib conjugates which, as shown earlier (***Section 5.3.1***) bind both the inactive and the active TKD conformations. As shown in ***Figure 5-24***, the mean lifetimes measured in cells labelled with Pel-Bod FL, without and with Pel-Bod 650, were remarkably consistent with those of Ner-Bod FL from the equivalent experiments (obtained from ***Figure 5-18***). This consistency between the type I and II fTKIs indicates that the number dimeric receptors with asymmetric configuration is small, if any.

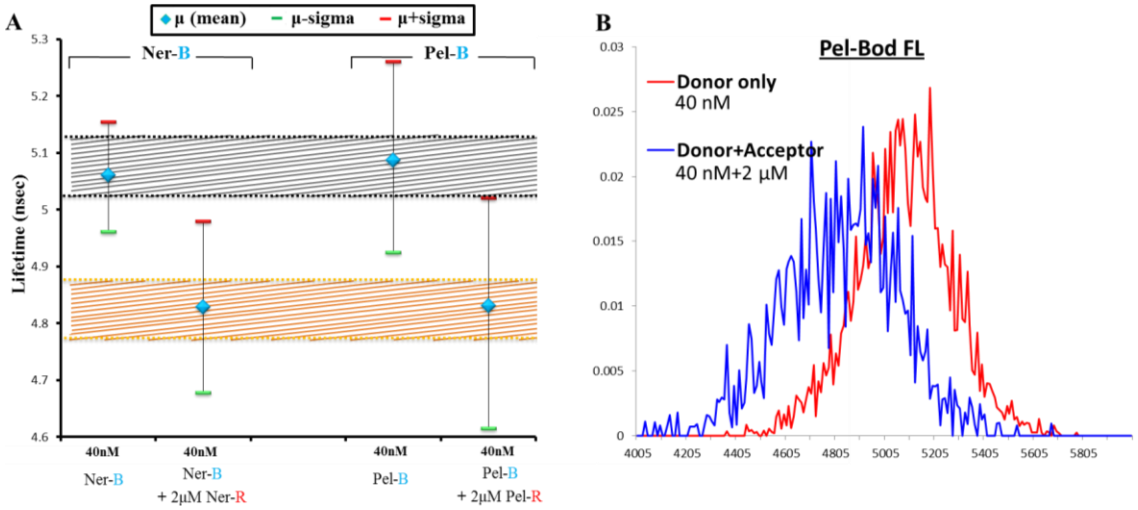


Figure 5-24: Comparison of donor fluorescent lifetimes in samples labelled with Neratinib or Pelitinib conjugates. CHO cells expressing wt EGFR were prepared as described in ***Figure 5-18***. Lifetimes of samples labelled with Pel-Bod FL (donor) without or with Pel-Bod 650 (acceptor). All experiments were performed in triplicate. (**A**) Comparison of Ner-Bod FL with Pel-Bod FL lifetimes in "Donor only" and in "Donor+Acceptor" samples. (**B**) Distribution of Pel-Bod FL lifetimes from (**A**).

In conclusion, the present FRET method constitutes the first, to the best of my knowledge, example of using small-molecule fluorescence probes for monitoring protein-protein interactions. The reduction in donor lifetimes of receptors labelled with both donor and acceptor molecules was small (***Figure 5-18*** and ***Figure 5-24***) and do not allow the drawing of strong conclusions. However, the fact that the lifetime of the high donor concentration samples (160-400 nM) remained unaffected by the addition of the acceptor (***Figure 5-21***) indicates that the observed FRET in ***Figure 5-18*** and ***Figure 5-24*** was due to bound fTKIs in close proximity. Previous independent studies assessing EGFR dimerization by FRET have shown that in the absence of EGF stimulation, or other treatments for inducing receptor dimerization, the receptor is mostly monomeric resulting in low FRET efficiencies[230,234,236]. Even though these results are consistent with the results presented herein and the weak FRET might be due to low dimerization, the lack of evidence of strong FRET signal from all conditions tested indicates that further studies are required. Experiments in EGF-stimulated cells or in highly dimeric cell lines will be performed in the future in order to assess the validity of the method.

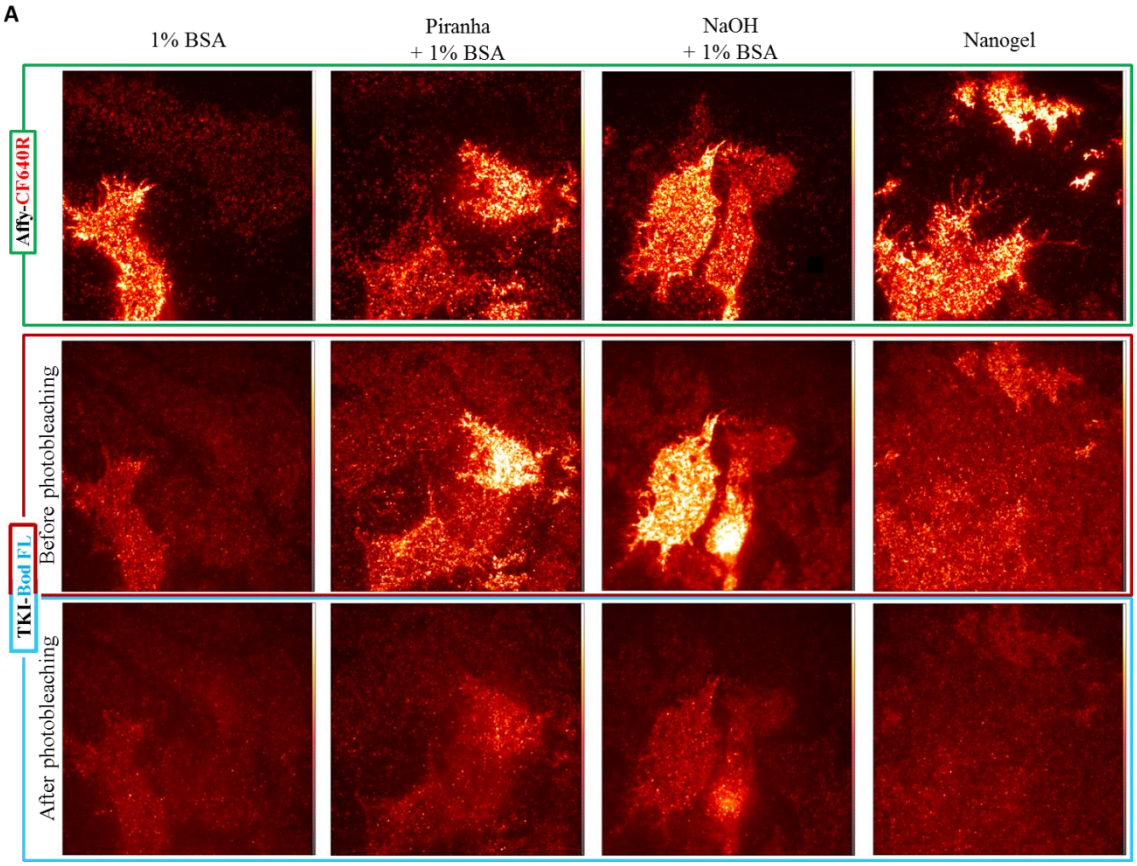
5.3.3 Utilization of fTKIs at Single-Molecule Level

TIRF microscopy and single-molecule techniques offer insight into the structure and the dynamics of biomolecules on the cell surface. My group's previous experience in single molecule methods (in both fixed and live cells)[70,71,104,161] led me to test the fTKIs in these techniques. As discussed in ***Section 1.3.2***, single-molecule methods such as SPT and localization-based methods report the distribution of values for a given property and allow the complete characterization of a system. In the last part of this chapter therefore I discuss the utilization and the challenges of using the small molecule probes at single-molecule level.

The only example available in literature where a small-molecule probe was used for single-molecule imaging was reported by Hern *et al*[246]. In order for the authors to study their target of interest (muscarinic receptors) they synthesized two fluorescently labelled telenzepine compounds (muscarinic antagonist) which specifically bind the receptor with picomolar affinity and employed them in single-particle tracking. The main difference between fTKI and these compounds is that the fTKIs bind to the intracellular domain of the receptor whereas telenzepine binds to the extracellular domain of

muscarinic receptors and therefore does not have to cross the plasma membrane. This last feature of telenzepine allowed the authors to incorporate moderate to highly negatively charged dyes (hydrophilic) which, as the work from Zanetti-Domingues *et al.*[161] has shown, are the best for single-molecule experiments due to the smallest possible introduction of artefacts.

Taking into consideration that fTKIs are labelled with hydrophobic dyes and previous experience of my group on the importance of surface (glass) passivation[104] for the non-specific binding of dyes (and hence the quality of the data), I sought to find the best treatment for the fTKIs. Bovine serum albumin (BSA) coating was employed in all four glass treatments tested, it has been demonstrated to reduce the non-specific binding of labelled protein molecules[104]. The four treatments tested were: BSA only, acidic (piranha) cleaning + BSA, basic (NaOH) cleaning + BSA and nanogel coating (PEG-BSA). The last one contains BSA and requires acidification and sylane treatment of the surface for efficient coating.



B

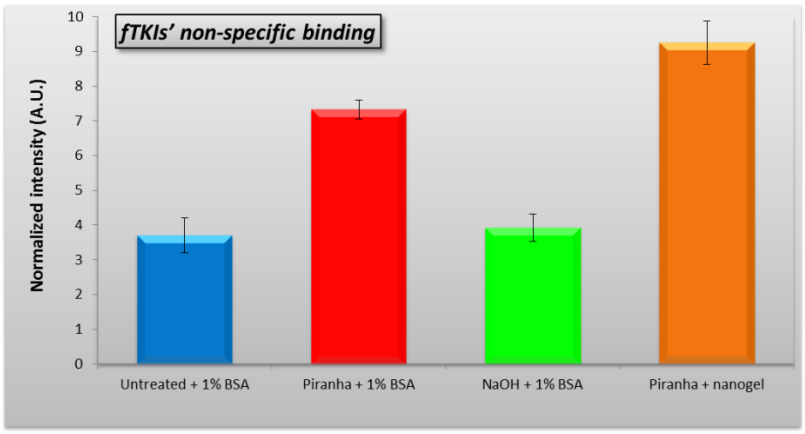


Figure 5-25: Non-specific binding of fTKIs at single-molecule level with different surface passivation treatments of glass surface. Cells expressing wt EGFR were seeded on glass-bottomed dishes passivated with four different treatments. After reaching the desired confluence, cell were labelled with Affibody and fTKI, simultaneously, for 1 h at 4^{o}C and then fixed. **A**) On the top row are images of the ECD (green box) whereas images in the middle and the bottom rows show the fTKIs labelling, before and after photobleaching. **B**) Quantification of non-specific binding in non-expressing cells.

Cells adhered and grew in all cases. I found that untreated (BSA only) dishes and dishes subjected to the basic treatment exhibited the lowest non-specific binding on the glass, with the former being slightly lower. Acidification of the glass (2nd and 4th columns, ***Figure 5-25A***) significantly increased the non-specific binding which remained high even after photobleaching (bottom row, ***Figure 5-25A***). Moreover the shape of the single-molecule spots in these two treatments was not as sharp as in the previous two. Despite the optimization, the signal-to-noise ratio in the fTKI channel was always poor compared to ECD labelling, however photobleaching can provide an additional strategy for improving the quality of the data. The disadvantage of the photobleaching approach is that the fraction of the immobilized molecules will not be monitored. An alternative method for imaging single molecules in fixed cells with improved signal-to-noise ratio, but still not as good as in ECD channel, was to label the receptor with fTKIs after fixation, as previously described for the quantitative analysis in the confocal mode. Thus, the incubation time could be significantly reduced as sufficient labelling for single-molecule studies could be achieved in 3-5 min.

Next I attempted to use fKTIs for live cell imaging and specifically for single-particle tracking (SPT). SPT allows one to study the dynamics of particles and how these are affected by different treatments, the confinement of the movement and the

frequency, as well as the duration, of particle interactions. As mentioned earlier, live cell imaging can be performed only by using irreversible fTKIs, therefore Pelitinib and Neratinib conjugates were used in the following experiments.

Live cell imaging requires cells to be maintained at 37°C and hence fTKI labelling must also take place at the same temperature. Even though the higher temperature negatively affected the non-specific binding, the SPT analysis algorithm[96] allows, in some extend, to distinguish the immobile (non-specific binding on the glass) from the mobile fraction of particles. Furthermore, during the optimization of the SPT protocol I found that a photobleaching-recovery strategy greatly improved the mobile-to-immobile fraction ratio (***Figure 5-26A***). In order to calculate the time required for single molecules to appear in the field of view after photobleaching, and before acquisition, I measured the diffusion coefficients of receptors after perturbation with unlabelled Neratinib (1 μM for 2 h) in both the basal and the activated state (***Figure 5-26B***). Given their diffusion coefficients and the channel size (80 x 30 μm) I estimated that 2 min are enough to allow molecules from surrounding areas to appear in the photobleached area.

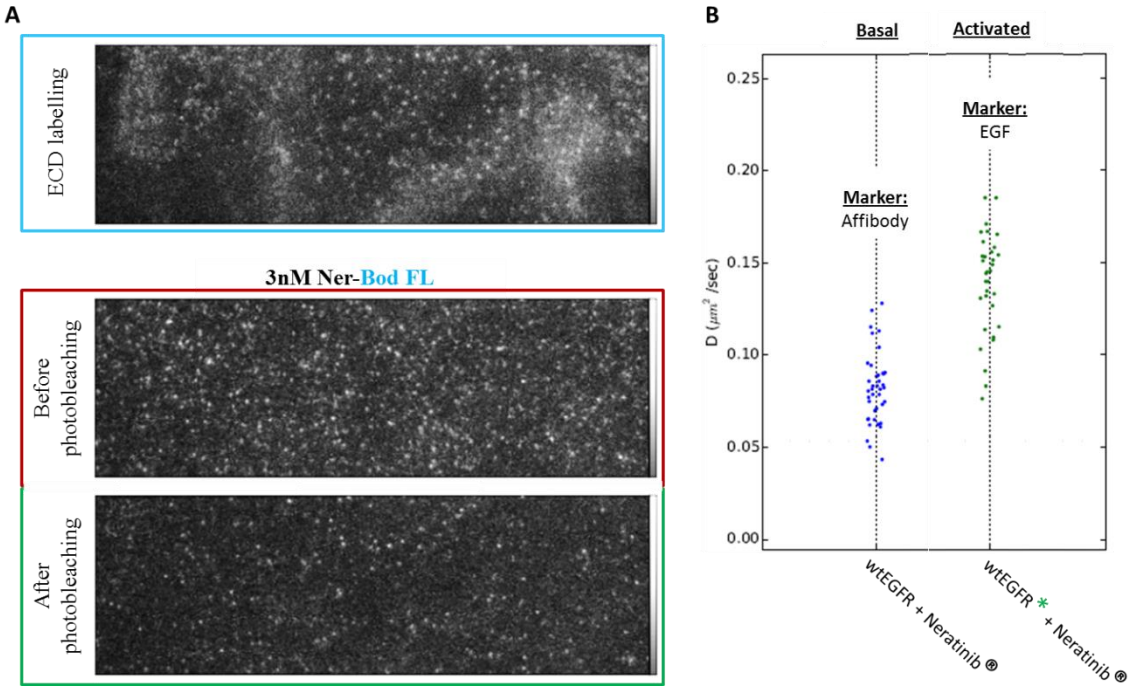


Figure 5-26: SPT images of cells labelled with ECD ligand and Ner-Bod FL and the diffusion coefficients of wt EGFR. A) Single-molecule images of live CHO cells expressing wt EGFRlabelled with ECD ligand (top image) and Ner-Bod FL before (middle image) and after photobleaching (bottom image). All experiments were performed in triplicate. **B**) Diffusion coefficient of wt EGFR in the basal and activated state after perturbation with 1 μM of Neratinib.

Having optimized the SPT protocol for the irreversible fTKIs I then designed an experiment for measuring the diffusion coefficient of the fTKI labelled receptors and comparing it with that of the ECD labelled receptors. For that I used a 2-colour SPT experiment whereby individual particles labelled with an ECD ligand and a fTKI can be tracked simultaneously. For ECD labelling I used either labelled EGFR Affibody for tracking the basal receptor or labelled EGF for the active state, both labelled with Alexa 488 or CF640R depending on the colour of the fTKI used. In order for me to be able to compare ECD with fTKI diffusion coefficients I needed to ensure that every receptor, in both conditions, has a ligand and an inhibitor bound to it. As single-molecule experiments require low labelling, for every labelling reagent cells were supplied with a saturating concentration of the equivalent unlabelled reagent. The protocol was as follow: cells were treated with 1 μM of inhibitor mixture (labelled and unlabelled) for 2 h, followed by labelling with ECD ligand mixture for 10 (Affibody) or 4 min (EGF) and then imaged at 37^{o}C.

The results (diffusion coefficients) from the SPT experiments are summarized in ***Figure 5-27***. Every coloured box within the plot highlights a set of experiments for every fTKI used (four in total). The first line in every box represents data obtained from fTKI channel in non-expressing cells treated as explained earlier. The second and the third lines show data from a two-colour experiment in cells expressing wt EGFR in the basal state. Affibody particles are shown in the 2nd whereas fTKI-labelled molecules in the 3rd. The fourth and the fifth lines depict the equivalent data for receptors in the activated state (EGF and fTKI respectively).

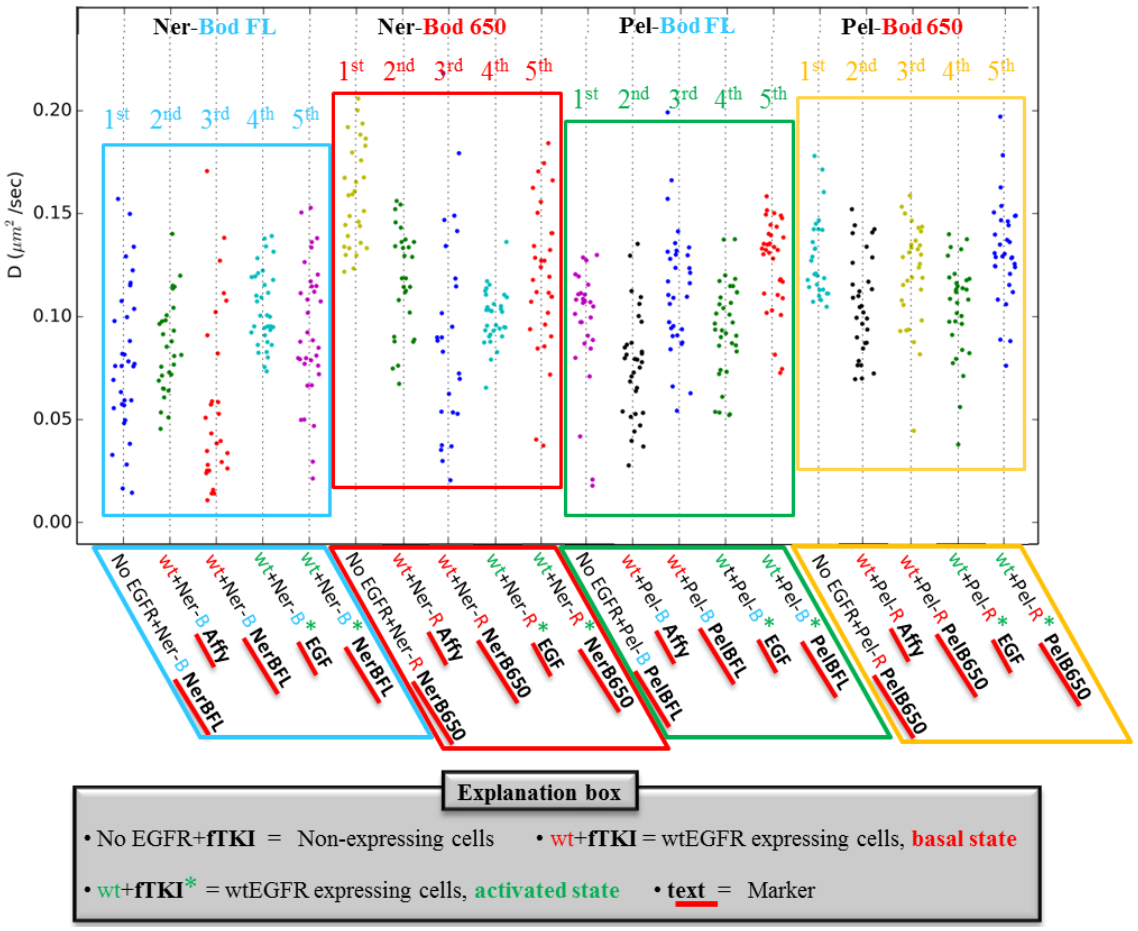


Figure 5-27: Irreversible fTKIs in Single-Particle Tracking. Particle diffusion rates (D) of receptors labelled simultaneously with a labelled ECD ligand (Affibody or EGF) and a reversible fTKI (Pelitinib or Neratinib) in CHO cells expressing wt EGFR. Cells were treated with a combination of both labelled and unlabelled reagents (ECD ligands and fTKIs) in order to ensure that all receptors have a ligand and an inhibitor bound to them. D was calculated from single particle tracks as described in ***Section 3.6.3.1. All experiments were performed in triplicate***. In each figure panel, one dot reports aggregated results from ~>10^3 tracks per field of view per dataset. Each point shows the estimate of D from a mean-square displacement (MSD) curve of the tracks within a field of view in each dataset.

Starting from the data in the 1st line, all fTKIs in the absence of the receptor produce a fraction of mobile particles which diffuse similarly to particles in expressing cells (3rd and 5th). The fact that these particles can be tracked by the SPT algorithm suggests that these molecules must be bound or trapped and not freely diffusing. Moreover, since TIRF allows imaging of samples within a depth of ~200 nm from the refection interface, fTKIs must be in the membrane region. During the evaluation in ***Section 5.2*** it was demonstrated that none of the fTKIs produce detectable membrane labelling in non-expressing cells. Even though this is not yet fully understood, I speculate that a small

fraction of fTKIs must be trapped within the membrane bilayer (most likely due to dyes properties) while they are in the process of permeating it.

Despite the overlap between the diffusion coefficients of fTKIs in non-expressing (1st lines) and expressing (3rd and 5th lines) cells, I noticed that the presence of the receptor on cells alters the dispersion of fTKIs data suggesting that the presence of the receptor might play a role. However, safe conclusions can be drawn and further optimization of the present method is required in order to minimize or overcome the penetration problems. This could be achieved by loading the cells with fTKIs using techniques such as microinjection[247], electroporation[248] or microfluidics[249] which has been shown that even allowed the delivery of quantum dots[250].

Overall in this chapter, the evaluation of fTKIs as imaging probes and their applications in different fluorescence imaging techniques were reported. All of the synthesised fluorescently labelled compounds exhibited low nanomolar affinity in biochemical assays. Apart from the TMR-labelled type I compounds (Gefitinib and Pelitinib), the rest fTKIs were cell permeable. As expected, reversible fTKIs (Gefitnib and Lapatinib) cannot be used for live cell imaging as there are being dissociated from their target. However, all compounds can be used in experiments for labelling the receptor after fixation. This post-fixation labelling method was used to study whether the type I TKIs are selective against the active EGFR conformation. The present results show for the first time that these inhibitors do have preference against the active conformation in cell environment. Furthermore, this conformationally selectivity can be suggests that fTKIs can be used as reporters for the activity of EGFR in cells.

Other applications reported herein include EGFR intradimer FLIM-FRET experiments as well as experiments at single-molecule level. However, both applications require further development and validation.

6 Architecture of the Epidermal Growth Factor Receptor on the cell surface: A Combinatorial Approach

6.1 Introduction

As shown in ***Section 5.3.1***, the combination of different fTKIs in quantitative binding studies can be used not only as a method to extract information regarding the mechanism of action of TKIs but also to understand the conformation of the kinase domain. I therefore sought to exploit this novel method in combination with methods previously reported by my group that can be used to probe the structure of the ECD to better understand both the conformational coupling of the ICD and the ECD and the architecture of EGFR on the plasma membrane.

The three methods I have employed to probe the ECD structure are: i) Distance of Closest Approach (DOCA) for measuring the distance between EGFR ligands and the membrane[136], ii) SPT for assessing the degree of non-monomeric complexes[137] and iii) FLImP for determining the geometry of the non-monomeric complexes[70,71,251]. I have focused my attempts on determining the architecture of the unliganded basal EGFR in its native membrane environment, complementary to recent work towards understanding the EGF-stimulated receptor[251].

6.2 The Conformation of the ECD of the Unliganded Receiver-Impaired Receptor is Extended

While it is generally accepted that the ECD conformation of the monomeric receptor is tethered[36] (***Figure 1-5A***), the ECD conformation adopted by the pre-formed inactive dimers is still under debate, as it was discussed in ***Section 1.2.2***. On one hand, negative stain electron microscopy (EM) images of EGFR from two independent studies, in which dimerization was driven either by stabilising the asymmetric kinase dimer interface with tyrosine kinase inhibitors[217] or by fusing the C-termini of the EGFR domain IV to dimerising proteins (Fc domain or the dimeric leucine zipper)[252], revealed two tethered ectodomains but with no well-defined interface between them. Moreover,

the tethered conformation of EGFR has also been suggested to be present even when heterodimerized with ErbB2, which is always extended, in the resting state[72].

On the other hand, seminal long-time molecular dynamics (MD) simulations of membrane embedded near full-length EGFR predicted that the ectodomains of ligand-free dimers adopt a back-to-back-like extended conformation[63]. The notion that the ectodomains of preformed ligand-free EGFR dimers adopt an extended conformation was also proposed by FRET measurements (DOCA) between the N-terminus of the ectodomain and the membrane[253,254]. Even though the distances reported (~8nm[253] and ~6.1nm[254]) are not the same in both studies (which may be attributed to different numbers of pre-formed dimers), both are still longer than the 4.5 nm N-terminal distance from the membrane predicted by MD simulations for the receptor monomer (tethered conformation)[255].

In order to investigate the conformation of the ECD of EGFR pre-formed dimers I have also employed the DOCA method developed by my group[136,251]. The present DOCA method is a lifetime FRET-based assay for measuring the distance between a fluorophore (donor) conjugated to EGFR specific ligands and a second fluorophore (acceptor) attached to a commercially available membrane probe (Vybrant DiD) on the outer leaflet of the plasma membrane (***Figure 6-1A***). As every donor molecule will most likely have more than one acceptor molecules in a distance shorter than 10 nm, the measured distance was determined from the variation of the FRET efficiency between fluorescent donors and acceptors measured as a function of acceptor surface density, based on Monte Carlo simulations[136].

In order to eliminate uncertainties that could be derived from measurements of a single ECD-membrane distance due to high monomer-dimer ratio (as previously discussed) or plasma membrane curvature[251], I studied the conformation by triangulating the separation from two different positions of the ECD to the membrane (***Figure 6-1B***). To achieve this, two fluorescent FRET donors were bound at domain I and III respectively in parallel experiments. One donor was a labelled nanobody (EgB4) that binds to domain I[220] (DI) and the second was a labelled anti-EGFR Affibody that binds domain III[256,257] (DIII). Both donor probes bind EGFR without inducing activation[70,220] and the dye I used was Alexa 488 (AF488). By using these two probes, if most EGFR ectodomains adopt a tethered conformation, the distance from DI

to the membrane will be shorter than that from DIII; if most ectodomains adopt an extended conformation, the opposite would be observed (***Figure 6-1B***).

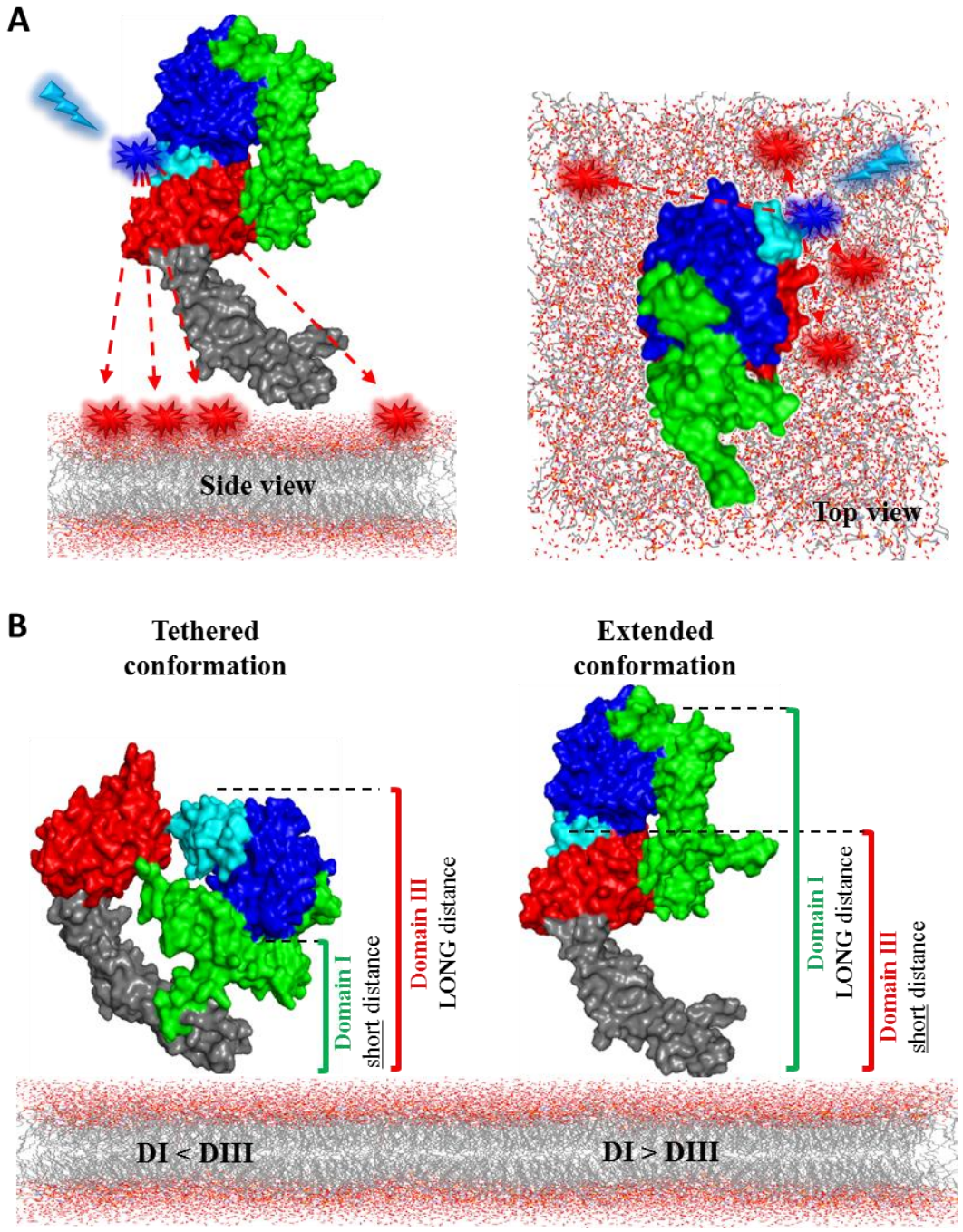


Figure 6-1: Schematic representation of DOCA principle and distances of ECD domains from the membrane in the different conformations. A) Labelling strategy for control FLIM-FRET based observations of labelled ligand-membrane FRET. A fluorescently labelled EGFR ligand (FRET donor) is bound specifically to the extracellular region of ErbB1 and the acceptor (Vybrant DiD) is free to diffuse in the outer leaflet of the plasma membrane. **B**) Illustration of distances of DI (green) and DIII (red) from the plasma membrane in the tethered (right) and the extended (left) ECD conformation. When the receptor adopts the tethered conformation DI distance is shorter than DIII whereas when the receptor is extended DI distance is longer.

I started the investigation using wt EGFR-expressing CHO cells, as previously, and performed parallel experiments in order to monitor the distances of both DI and DIII. Cells were first treated with the membrane probe for 15 min at 37 ^{o}C followed by labelling with either EgB4 or Affibody, both conjugated to AF488, for 1 h at 4^{o}C and finally fixed. I acquired both confocal images of DiD signal intensity (acceptor channel) and lifetime images of ECD probes.

The average distance of the ECD probes to the membrane (***Figure 6-2C***) was calculated from the plots of the FRET efficiency as a function of acceptor density (***Figure 6-2 A&B***) using the best fits of Monte Carlo simulation as previously described[136]. Even though the distances of both DI and DIII are very similar (6.9 ± 0.6 nm and 6.6 ± 0.6 nm respectively), the fact that DI distance is consistent with a previous study suggesting the extended conformation[254] and that it is also longer than the DI distance of tethered monomers predicted by MD simulations[255] suggest a contribution of some receptors which adopt the extended conformation.

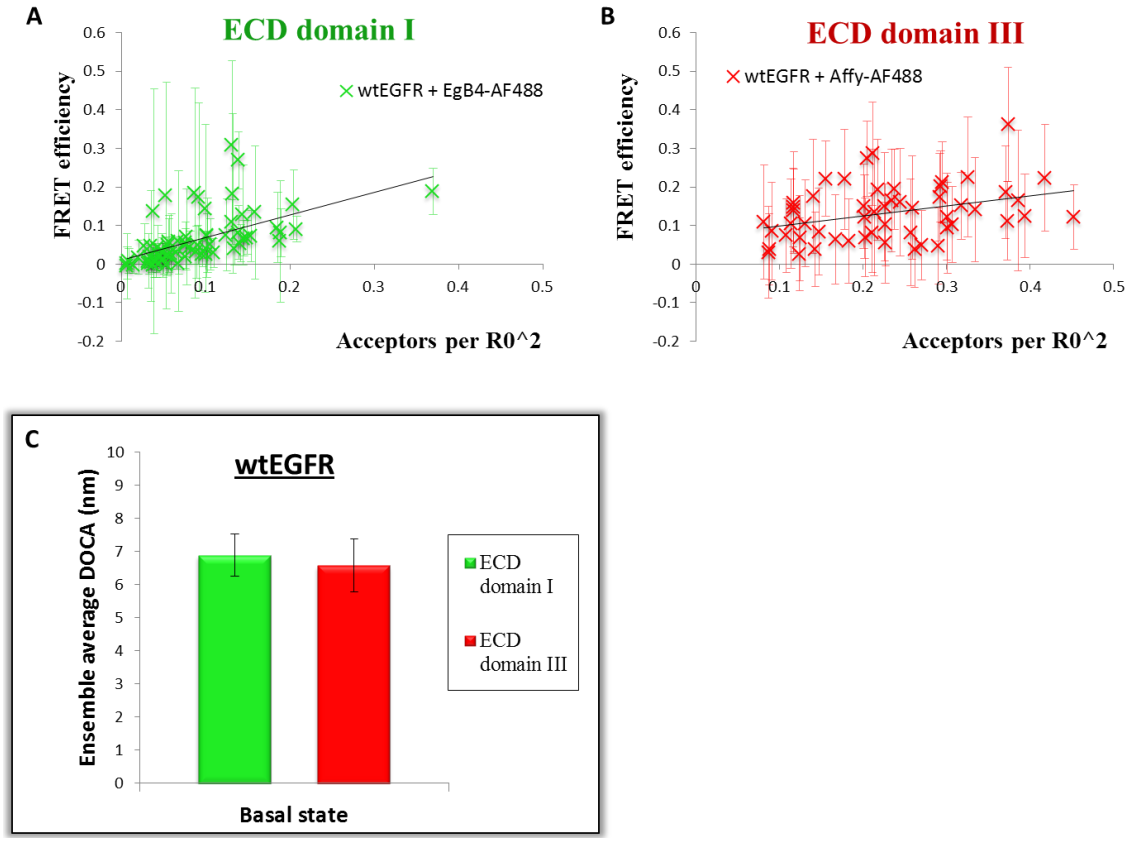


Figure 6-2: Determination of the distance of closest approach (DOCA) of EGFR ligands to the cell surface. A&B) Plots of FRET efficiency as a function of acceptor density measured in acceptor (DiD)-loaded CHO cells expressing wt EGFR. The receptors were labelled with either EgB4-AF488 (**A**) or Affy-AF488 (**B**). All experiments were performed in triplicate. The best fits of Monte Carlo simulation results to the data are shown. The errors from

uncertainty in the Monte Carlo model fitting were typically 10-20% of the calculated distance values as previously described[136]. **C**) Distances between EGFR-bound ligands and the membrane, derived from FRET measurements shown in ***Figure 6-2 A&B***.

An explanation for the fact that DI and DIII distances from the membrane are almost indistinguishable suggests a mixed population of both tethered and extended ectodomains in equilibrium. Furthermore, the results suggest that the extended population consists a considerable proportion of the total population, given that monomers adopt the tethered conformation[36] and that the DOCA method does not allow distinguishing between monomeric and non-monomeric receptors. The presence of a mixed population of ECD conformations is further supported by the fTKI results (***Section 5.3.1***) which showed the presence of both active (binding of f-Gef) and inactive (binding of both type II fTKIs) TKD conformers in the basal state of the wt receptor. However the fact that type II fTKIs exhibited higher binding of in the basal receptor compared to type I suggested that the dominant species in the basal state is constituted of receptors with inactive TKDs (monomers and/or symmetric dimers). Collectively, the results from these two methods led me to speculate that an unliganded extended ectodomain conformation is coupled to an inactive TKD conformation (either monomers or symmetric dimers). Although the monomer-dimer ratio in the cell line used is not known, the conformational coupling could be ascribed to dimeric receptors if the general assumption that EGFR monomers adopt an ECD tethered conformation (and of course an inactive TKD) is valid[36]. It is worth noting that this conformational coupling between unliganded extended ECD dimers and symmetric TKDs dimers has also been predicted by *in silico* studies[63].

In order to further investigate the hypothesis for the conformation coupling of the inactive unliganded receptor I sought to determine the conformation of a receptor that cannot be activated in order to eliminate the receptor population responsible for f-Gef binding (and for the basal phosphorylation, ***Figure 5-13***). For that I employed a constitutively inactive EGFR mutant (L680N)[49] with a single-point mutation on the N-lobe of the TKD which disrupts the asymmetric dimer formation but not the formation of the symmetric dimers (***Figure 6-3A***). If the above conclusion is true and if inside-out signal regulation occurs[217], the ECD DI distance of such a mutant would be higher than 6.8 nm as the ratio of tethered-to-extended ECD would be decreased. Moreover, f-Gef

binding is expected to be decreased as active kinases would be absent (f-Gef binding was shown in ***Section 5.3.1*** to be proportional to active/phosphorylated kinases).

The experiments were conducted in transiently transfected CHO cells expressing L680N EGFR. I first determined the distances of both DI and DIII using the DOCA method as described for wt EGFR. As expected, not only the distance of DI from the membrane increased by nearly 2 nm compared to wt EGFR (from 6.89 ± 0.6 nm to 8.67 ± 0.64 nm) but also the DI distance was distinctly longer than the DIII distance, which only marginally increased (6.97 ± 0.8 nm, ***Figure 6-3B***). These results suggest that the predominant ECD conformation of the basal (unliganded) asymmetric dimer-impaired receptor is extended. Furthermore, the fact that exclusion of one, out of three possible (***Figure 6-3A***), TKD configurations (the asymmetric dimer) resulted in the "fading" or disappearance of an ECD configuration supports the inside-out regulation of ErbB receptors.

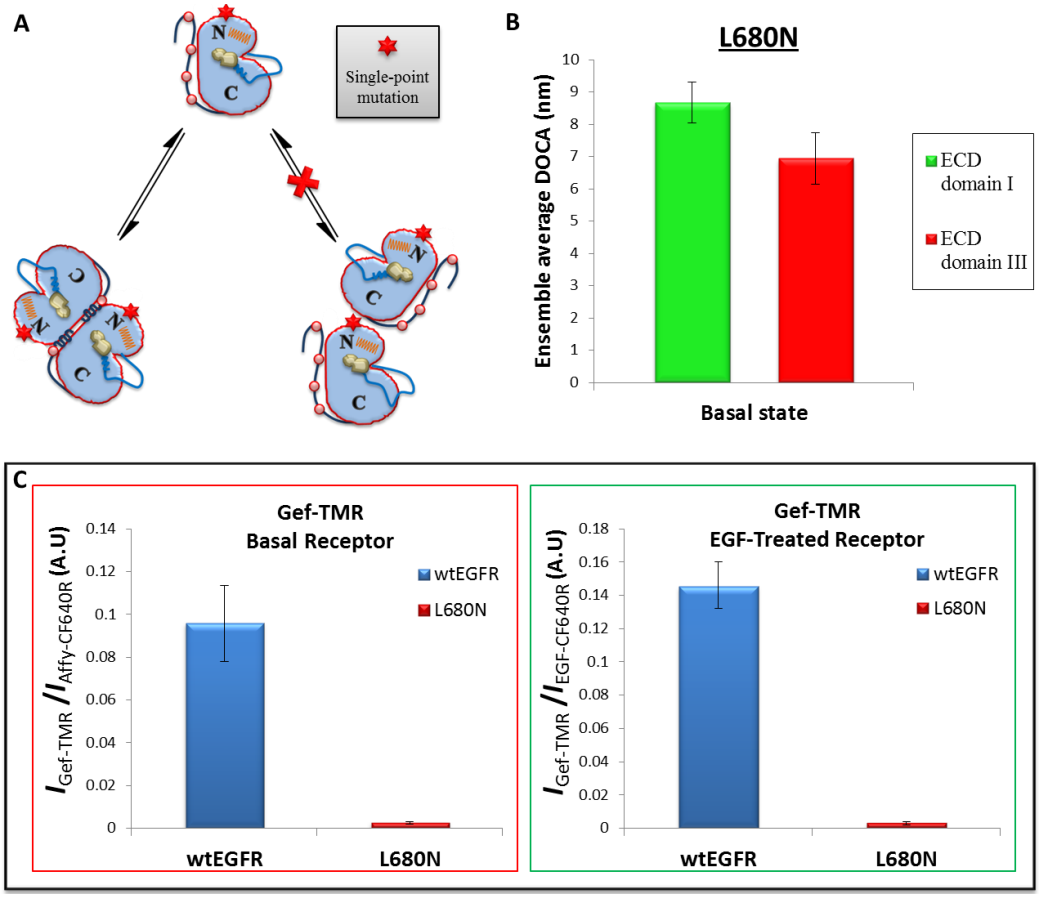


Figure 6-3: Determination of the ECD and TKD conformation of a constitutively inactive EGFR mutant. A) Cartoon representation of feasible conformations of L680N EGFR. **B**) Determination of DI and DIII distances in the basal state of L680N. **C**) Gef-TMR binding in L680N-expressing CHO cells (red column bars) in both basal (left) and

EGF-treated (right) state and comparison with its binding in wt EGFR cells as a control (blue column bars) . All experiments were performed in triplicate.

For determining the TKD conformation of L680N I employed the method described in ***Section 5.3.1***. The results from the experiments using f-Gef are shown in ***Figure 6-3C***. In this, as well as in all the following plots, the results from experiments in wt EGFR-expressing cells are also included as a control. As expected, disruption of the asymmetric dimer interface completely abolished Gefitinib binding in the basal state (***Figure 6-3C***, left plot). In order to ensure that EGF binding is unable to alter the TKD conformation I also performed the experiment in EGF-treated cells and as I expected f-Gef binding remained unaffected (***Figure 6-3C***, right plot).

As opposed to f-Gef, I expected that the other three fTKIs (f-Pel, f-Lap and f-Ner) would label L680N with similar efficiency as the basal wt EGFR was labelled since all three of them were shown to bind inactive TKDs (***Section 5.3.1***). However this was not the case: compared to wt EGFR-expressing cells a ~5.5-fold decrease was obtained in all three cases for both basal and EGF-treated L680N EGFR expressing cells (***Figure 6-4 A-C***). Furthermore the fact that the binding of all three fTKIs was equally decreased regardless of their generation (reversible or irreversible) and/or type (I or II), but also irrespective of ligand binding, shows the reduced ability (compared to wt EGFR) of all inhibitors to associate (K_{on}) with the ATP pocket. Despite this decrease due to altered K_{on} of these fTKIs (discussed below), for all three fTKIs their fluorescence signal on the membrane of the L680N EGFR mutant was significantly higher compared to f-Gef indicating the presence of only inactive kinases (also consistent with western blot analysis reported by Zhang *et al.*[49]). This last result clearly suggests that the unliganded extended ECD (DOCA results of L680N EGFR) is conformationally coupled with the inactive TKD (fTKIs results of L680N EGFR).

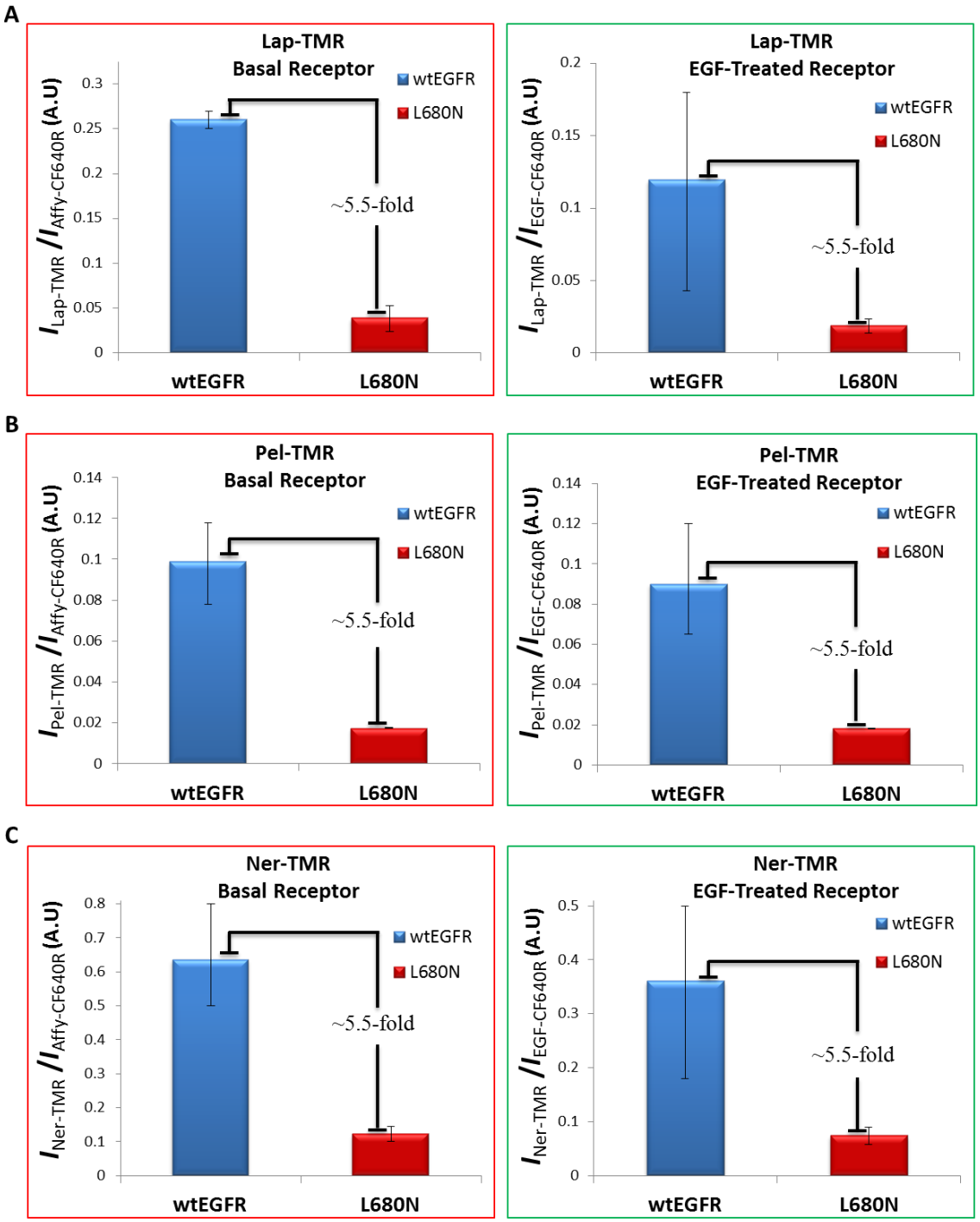


Figure 6-4: Type I irreversible and type II (reversible and irreversible) fTKIs binding in L680N EGFR-expressing cells in basal and EGF-treated cells. Lap-TMR (**A**), Pel-TMR (**B**) and Ner-TMR (**C**) binding in L680N EGFR-expressing CHO cells (red column bars) in both basal (left, red outline) and EGF-treated (right, green outline) state and comparison with their binding in wt EGFR-expressing cells as a control (blue column bars). Data were analyzed as discussed in ***Section 5.3.1***. All experiments were performed in triplicate.

As mentioned in the previous paragraph, the equal reduction in the binding of f-Pel, f-Lap and f-Ner (which all bind to the inactive conformation) in L680N EGFR

expressing cells is due to their anibility, at least compared to wt receptor, to associate with their target. This reduction could be either due to a significantly altered ATP pocket conformation or due to inaccessible ATP clefts.

An alteration of the shape of the ATP-site, which would result in reduced binding affinity of TKIs, would mean that the L680N mutant adopts a conformation distinct from both inactive and active wt EGFR conformation. However, crystal structures of Lapatinib-bound (inactive conformation) wt EGFR (***Figure 5-11A***), activator-impaired (V924R or V948R) EGFR (***Figure 5-11C***) and receiver-impaired (I682Q, similar to L680N) EGFR[258] have shown that the inactive canonical ATP conformation is preserved in all three of them. Nevertheless, in order to further explore the binding of Pelitinib and the type II fTKIs against the inactive TKD conformation I employed a kinase-dead EGFR mutant (K721A)[49]. Lys[721] is a key residue located within the ATP pocket which stabilizes the active conformation of the TKD, after allosteric activation, by interacting with a glutamine residue located on the αC-helix which in turn keeps the αC-helix in the "in" position[53,259]. It has been shown that mutating this residue to alanine or methionine not only impairs the ability of the receptor to catalyse the transfer of the ATP phosphate group to tyrosine residues but also prevents the formation of the active TKD conformation in the receiver kinase[53] (***Figure 6-5***). Other than these two important differences, K721A mutant should behave as the wt receptor.

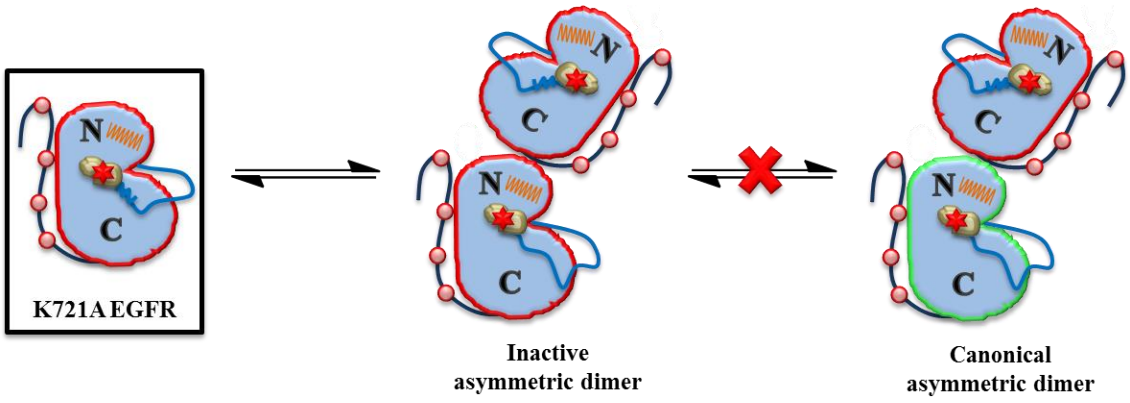


Figure 6-5: Permitted TKD receiver conformation in the asymmetric dimer configuration of the K721A EGFR mutant. Mutation of the catalytic lysine residue located within the ATP pocket allows the formation of an asymmetric dimer where both kinases are in the inactive (red outline) conformation (middle dimer) as opposed to the canonical dimer where the receiver adopts the active (green outline) conformation.

Since the TKD conformation of K721A EGFR is always inactive and if the inactive ATP conformation is not responsible for the reduced fTKI binding, as I previously

discussed, then f-Pel and both type II fTKIs should bind the K721A with similar affinity as the wt EGFR. The results from the experiments using the K721A EGFR mutant are shown below (***Figure 6-6***).

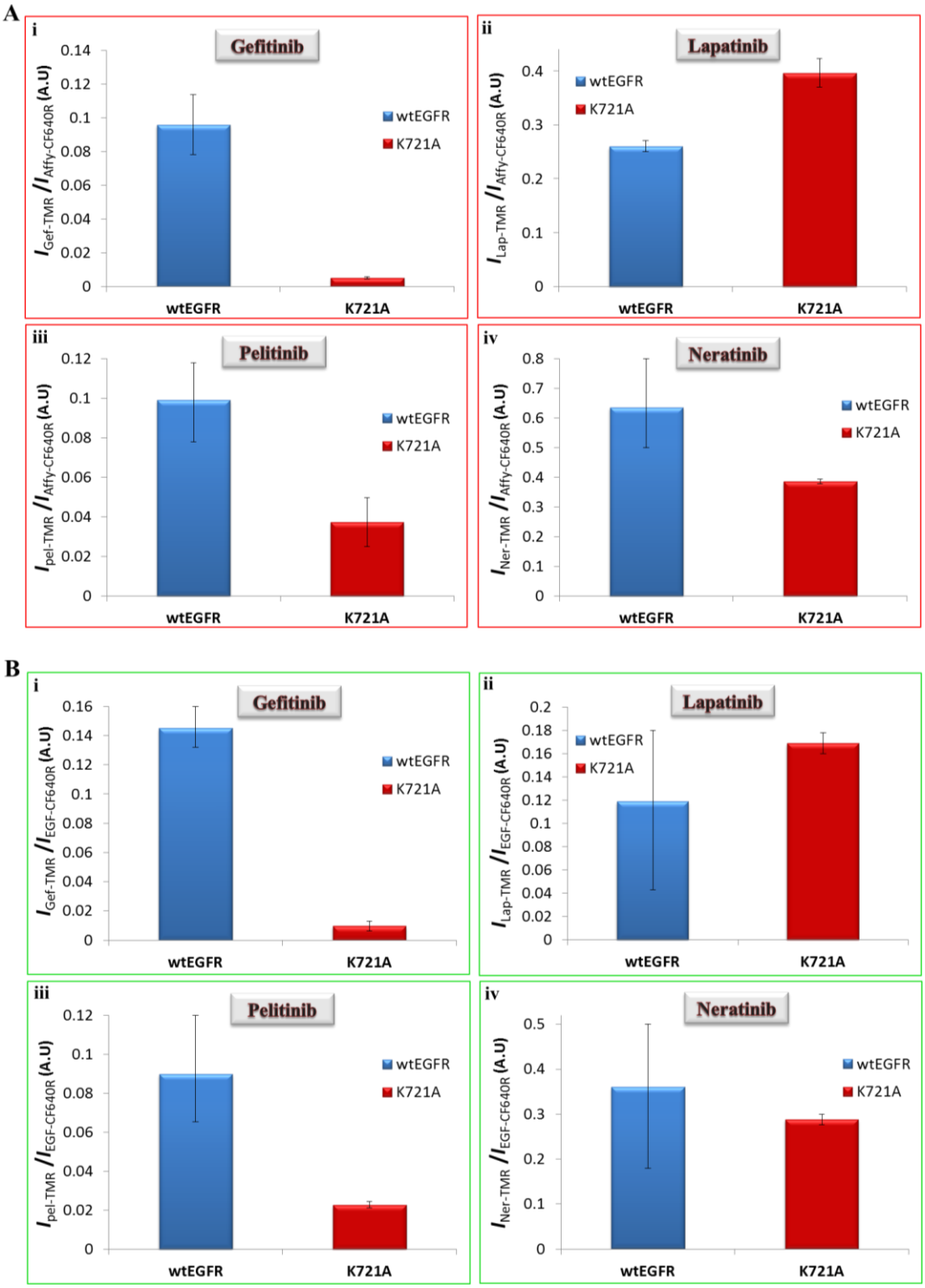


Figure 6-6: fTKIs binding in K721A-expressing cells in basal (A) and EGF-treated cells (B). fTKIs binding in K721A EGFR-expressing CHO cells (red column bars) in both basal (**A**) and EGF-treated (**B**) state and comparison with their binding in wt EGFR-expressing cells as a control (blue column bars). In both (**A**) and (**B**), Gef-TMR data is

shown in (**i**), Lap-TMR in (**ii**), Pel-TMR in (**iii**) and Ner-TMR in (**iv**). Data were analyzed as discussed in ***Section 5.3.1***. All experiments were performed in triplicate.

Starting with Gefitinib, its binding to K721A was abolished (as in the case of L680N) as a result of only inactive kinases in both basal and EGF-treated cells and despite the presence of asymmetric dimers (***Figure 6-6 A&B, i***). The binding to K721A was only marginally higher than to L680N, 0.0051 and 0.0027 A.U. respectively. On the other hand, Lapatinib membrane signal in K721A-expressing cells was higher compared to wt EGFR, in contrast to binding in L680N cells. This increase is consistent with the higher number of inactive kinases in the K721A cells after the "replacement" of the active kinases from the basal phosphorylated population of the wt EGFR by inactive TKDs (***Figure 6-6 A&B, ii***). Even though I would expect, based on the results obtained with Lapatinib, the same increase in the binding of both irreversible fTKIs to K721A, both fTKIs exhibited lower affinity, but nevertheless higher than in L680N. The answer to this unanticipated result may lay in the structural differences between Lapatinib and the irreversible TKIs. Structurally, both Lapatinib and Neratinib (but also the smaller Pelitinib) are very similar and show a common binding mode (as crystallographic data have shown[47,51,260]). Therefore the reason for the decreased affinity of the latter (and Pelitinib) could not be due to any structural restriction related to their size (***Figure 6-7***). Besides, the fact that Lapatinib is able to bind the receptor is a strong indication of an inactive ATP conformation.

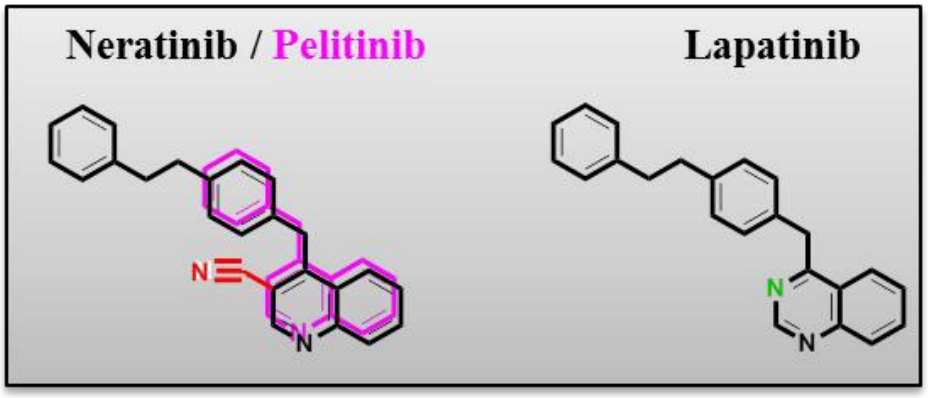


Figure 6-7: Structural similarities and differences between Neratinib/Pelitinib and Lapatinib. Lapatinib is not only structurally very similar to Neratinib and to the smaller Pelitinib (purple superimposed structure) which lacks the additional benzyl ring characteristic of type II TKIs, but it also shares the same binding mode as crystal structures have shown. The only, yet essential, difference lies on one of the features responsible for a H-bond interaction of all three TKIs with the ATP-site (red and green coloured structural parts).

As discussed in ***Section 4.3***, both Pelitinib and Neratinib have been developed by Wyeth and share a feature which is distinct from the other EGFR anilinoquinazoline inhibitors such as Gefitinib, Lapatinib, Erlotinib and *etc*. This common feature is an "extended" nitrogen atom on the cyano-group (red colour, left-hand side molecules in ***Figure 6-7***) which directly interacts with a threonine residue, as opposed to Lapatinib (and other inhibitors) which requires the presence of a water molecule to bridge the interaction between the nitrogen atom on the quinazoline ring (green colour, right-hand side molecule in ***Figure 6-7***) and the same threonine. This is the only notable difference between Lapatinib and Neratinib/Pelitinib as this feature is involved in one, out of only two, H-bond interactions. Taking into account that the mutated residue in the K721A is located within the pocket and that there is no difference in the binding between these molecules in the wt receptor, I concluded that the reduced affinity of both Neratinib and Pelitinib is due to steric hindrance from the presence of the "extended" nitrogen of the cyano-group. The fact that Lapatinib binding was increased in K721A indicates that the reduced binding against L680N EGFR shown earlier (***Figure 6-4A***) cannot be due to altered ATP conformation of the receptor in the inactive state.

Inaccessibility of the ATP-sites in the symmetric dimer configuration has been proposed by Arkhipov *et al.*[63] using long-timescale MD simulations (***Figure 6-8A***). They found that some basic residues on the TKD surface, which are exposed in the symmetric but not in the asymmetric dimer, electrostatically interact with the inner leaflet of the plasma membrane and that this interaction results in sequestered ATP clefts. The interaction of the inactive TKD with the membrane has also been proposed by other studies suggesting a model where activation involves release of the kinase domains from the membrane[54,261,262]. Such an occlusion of the ATP clefts could explain why the K_{on} of several different TKIs has changed.

Blockage of the ATP binding site could also result from intermolecular interaction between TKDs. Such an interaction was found after examination of the crystal structures of both the activator-impaired (V924R) and the receiver-impaired (I682Q) EGFR, resolved by Kovacs *et a*l[258]. Apart from the canonical symmetric dimer (***Figure 6-8B***), the crystal lattice of both mutants contained an alternative configuration where the surface with the ATP-site opening of one kinase is blocked by the backbone of the other kinase (***Figure 6-8C***). Even though such an interaction would be consistent with

the observed reduced fTKI binding, it is unclear whether this configuration is biologically relevant (it could be related to higher order oligomers formation, for instance), as it is not commented upon by the authors.

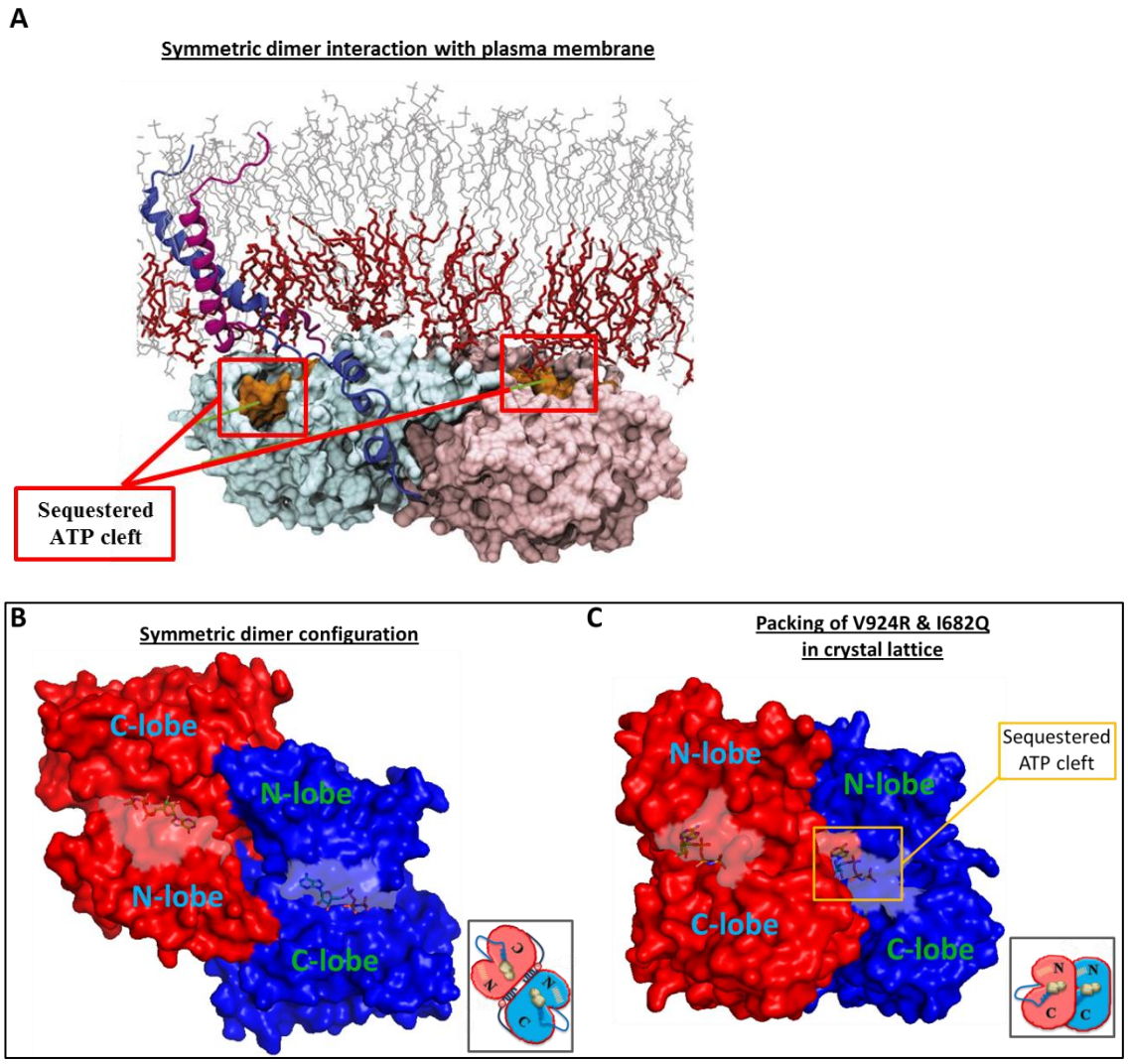


Figure 6-8: Accessibility of ATP clefts of symmetric KD dimers. A) Illustration of electrostatic interaction beween symmetric dimer and plasma membrane **(**Adapted from Arkhipov *et al*., Cell, 2013**). B**) Configuration of symmetric dimer found in crystal structures with PDB name 3GT8 and 5CNN. **C**) Additional interacting species with back-to-front configuration found in the crystal lattice of both aforementioned crystal structures.

Irrespective of the cause for the sequestered ATP clefts in L680N EGFR expressing cells, the fTKIs binding results revealed that the predominant TKD conformation of a receptor with disrupted asymmetric dimer interface (L680N) is much less favoured in the basal "inactive" wt receptor. This observation is also consistent with the results reported by Mi *et al.*[65]. EM images showed that mutation (T669D/S671D) in the JM segment which latches the C-lobe of the activator (similarly to L680N which is located

closer to the C-terminal end of the JM than T669D/S671D) in the asymmetric dimer interface promotes symmetric dimerization, a conformation adopted by only ~1/3 of the wt EGFR population. If this conformational preference of the receptor is representative of the conformations that occur on the cell surface, then my results, in agreement with Mi *et al.*, suggest that symmetric dimerization is less favoured when the asymmetric dimer interface is intact. Besides, wt EGFR (without perturbation by TKIs) is always crystallized in the asymmetric dimer configuration and symmetric dimer crystals can only be obtained after introduction of either the N-lobe or the C-lobe mutation. The only ErbB member that has been crystallized in the inactive conformation while being in the wt form is ErbB4, but even in that case the resolved crystal was monomeric[51].

All the above do not necessarily mean that wt EGFR predominantly forms asymmetric dimers in the absence of ligand unless asymmetric dimers such as the K721M EGFR (mentioned earlier, ***Figure 6-5***) and the wt ErbB2 (reported by Aertgeerts *et al.*[55]), where both kinases are in the inactive conformation, do exist for wt EGFR too. This would contradict both the high binding of type II fTKIs (which only bind to the inactive kinases) in the basal wt EGFR state (***Figure 5-16 A&B***) and the low basal phosphorylation (***Figure 5-13***). The high binding of type II fTKIs could be due to binding to inactive monomers, to the "inactive" activator on the asymmetric dimers (discussed later), or to the slightly tilted symmetric dimers (with partially accessible ATP clefts) recently suggested by Bragin *et al.*[75] as a result to their coupling with a dimeric "inactive" JM structure which is different from what it was previously thought (***Figure 1-8C***).

In conclusion, I have found that the extended unliganded ECD conformation is favoured in the inactive (and presumably non-monomeric) population of the basal receptor as a result of its conformational coupling with symmetric dimers. The fact that the extended unliganded conformation is an inactive form of EGFR has also been proposed by Walker *et al.*[263]. The authors reported that disruption of the auto-inhibited (tethered) ECD conformation, and subsequent promotion of the extended conformation in the absence of ligand, by mutation did not result in activation and that this conformation is not sufficient for receptor activation as ligand binding is still required. Furthermore, both DOCA and fTKI binding experiments indicated that this inactive population is not the predominant one in the basal wt receptor. Lastly, the asymmetric kinase dimers are present in the basal state in wt EGFR-expressing cells (as both the

basal phosphorylation and Gefitinib binding indicated, ***Figure 5-13*** and ***Figure 5-12*** respectively) and since the ECD unliganded extended conformation can only be coupled to one TKD conformation, otherwise regulation of EGFR activation in healthy cells would not be possible, by elimination the ECD conformation that remains is surprisingly the tethered.

6.2.1 The Unliganded Extended ECD Conformation Promotes Non-Monomeric Complexes

I speculated that if the basal unliganded L680N receptor adopts an extended ECD conformation in which the same ECD dimeric interface of the liganded receptor, known to promote dimerization, is exposed, it should also promote higher number of non-monomeric complexes compared to basal wt EGFR. In order to investigate this I sought to exploit two-colour SPT to compare the number and the duration of the complexes formed.

Two-colour SPT experiments combined with advanced data analysis algorithms allows one to determine the relative (absolute quantification is not possible) frequency of pairwise interactions (colocalization) between two single particles (tracks) in two channels as well as the duration (T_{on}) of every interaction[96,137,251]. The basic principle of the colocalization experiments is shown below (***Figure 6-9***). Briefly, live cells were labelled with two different fluorescent dyes (always Alexa-488 and CF640R[161]) conjugated to either anti-EGFR Affibody or EGF. After acquisition of two-colour TIRF images, separated by a beam splitter[73], over 30 seconds for each area, the two channels were registered and single particle tracks were extracted as previously described[96]. In order for an event to be regarded as a colocalization event two tracks in different colour channels need to spend at least 3 frames (150 ms) moving together while remaining within a distance of each other of less than one pixel. The duration of individual events (until they move apart again) can be also calculated.

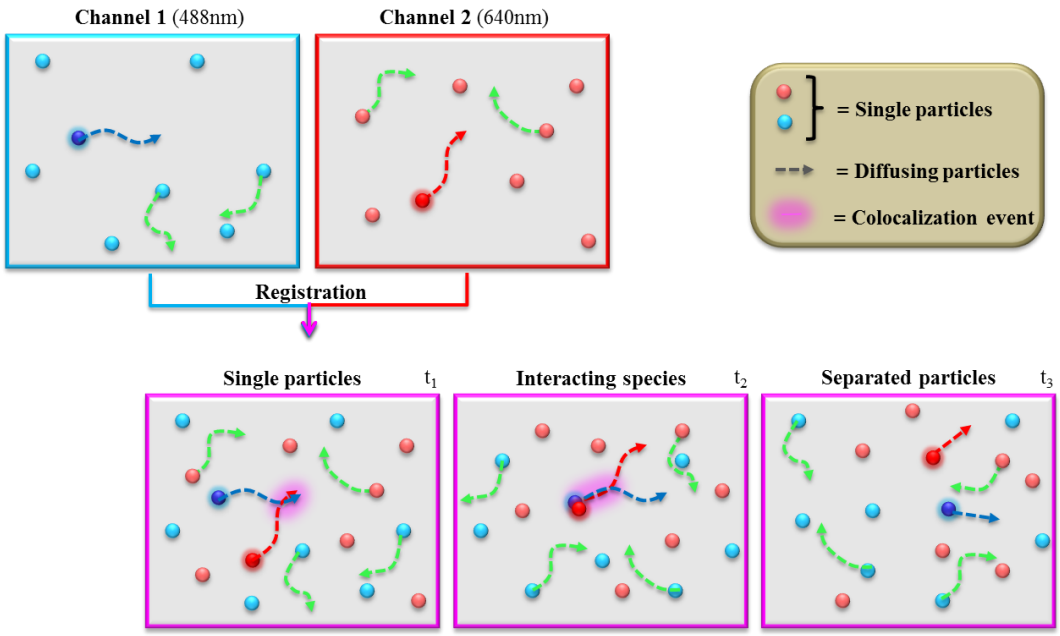


Figure 6-9: Two colour single-particle tracking method for determining colocalization frequencies and duration (T_{on}) of colocalized events. Cells expressing the protein of interest were labelled with two different fluorescent dyes and data were acquired. The two channels were then registered and single particle tracks were extracted as previously described[96]. Tracks from different channels which colocalize and move together can be extracted from which parameters such as colocalization frequency and T_{on} can be calculated.

For two-colour SPT experiments, cells were maintained at 37°C at all times during cell preparation and data acquisition and treated with pre-heated labelling solution. In order to determine whether the exposure of the ECD dimer interface in basal L680N-expressing cells promotes non-monomeric complexes I conducted two control experiments. In the first I measured the colocalization frequency and T_{on} of basal wt EGFR (labelled with anti-EGFR affibody), while in the second I repeated the experiment in EGF-stimulated (8nM) wt EGFR cells (labelled EGF conjugates were used as markers). For the basal L680N EGFR experiment, the same conditions as in basal wt EGFR were employed. An overview of the distributions of pairwise colocalisation frequencies and T_{on} for all receptor pairs is provided in ***Figure 6-10***.

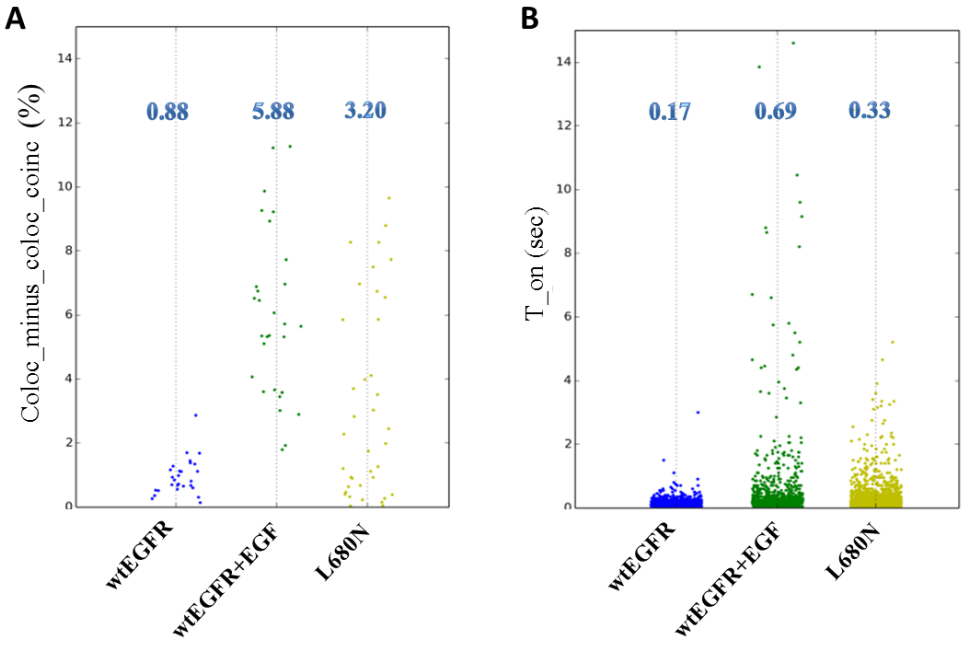


Figure 6-10: Fraction of colocalized tracks (A) and duration of the colocalization events (B) of basal and EGF-treated wt EGFR cells and basal L680N EGFR cells. Cells expressing with wt EGFR or L680N EGFR were labelled or labelled/stimulated with Alexa-488 and CF640R conjugates of anti-EGFR-Affibody or EGF respectively. **A**) Fractions of tracks in which the features in two tracks in different colour channels spent at least 3 frames (150 ms) moving together while remaining within a distance of each other of less than one pixel. B) Duration of the colocalisation events from the time a feature in one channel begins to move together within a pixel of a feature in the other channel until then they move apart. All experiments were performed in triplicate.

Starting from the control experiments, the data showed that while the wt receptor in the basal state (negative control, blue dots) forms only a small number of non-monomer complexes (0.9±0.6), EGF stimulation (positive control, green dots) greatly increased the fraction of colocalized tracks (~6.5-fold) between the two channels (***Figure 6-10A***). Disruption of the asymmetric dimer interface (L680N mutation), with concomitant promotion of the extended ECD conformation, in the basal state however resulted in a ~3.5-fold increase in the colocalization frequency compared to the basal wt EGFR. I also found that L680N EGFR complexes are interacting with each other more strongly than the unliganded wt EGFR but less strongly than the EGF-bound wt EGFR complexes ($T_{on\text{wt EGFR-EGF}} \approx 2 \times T_{on\text{L680N}}$). The latter might be explained by the fact that dimerization/oligomerization in L680N EGFR (TKD mutation) is driven only by the TKDs with a less favoured (still more favourable than the basal wt receptor), ECD conformation (unliganded extended conformation) as predicted by MD simulation[63]. On the other hand, EGF binding induces conformational changes that favour the formation of complexes which are stabilized from both the outside (crystallographic extended

EGF-bound dimer) and the inside (crystallographic asymmetric dimer). Therefore, these results support the previous assumption that stabilization of an unliganded extended ECD conformation through an inside-out regulation is accompanied by promotion of non-monomer complexes. Furthermore, the fact that L680N EGFR forms a higher number of non-monomers also agrees with the notion that the dominant TKD conformation of L680N EGFR (which is less favoured in the basal wt EGFR-expressing cells) is responsible for the decreased fTKIs affinities is the symmetric dimer.

6.3 Geometry of EGFR on the Plasma Membrane

Even though the DOCA and fTKI data in ***Section 6.2*** suggested that unliganded EGFR dimers with an extended ECD and a symmetric TKD conformations are not the dominant species in the wt receptor (as in L680N EGFR expressing cells), they do not give insight into the existence and the proportion of the different dimeric populations of the receptor (the possibility for the presence of a tethered ECD dimer was discussed in ***Section 6.2***). In order to be able to detect the two (extended and tethered) dimeric conformations I require a method which is capable of achieving sub-6 nm resolution. Such a method, namely Fluorophore Localization Imaging with Photobleaching (FLImP), has been previously published by my group for reporting lateral separations in the 0-60 nm range between identical fluorophores bound to the ECD[70,71,264]. I therefore employed this method not only for distinguishing between the two dimeric conformations but also for better understanding EGFR's architecture on cells.

FLImP allows the measurement of discrete pairwise separations between fluorophore-conjugated ligands bound to EGFR complexes. Unlike the single molecule analysis by Kuriyan and colleagues[265] where the fluorophores were attached to EGFR intracellularly, here the fluorophores are positioned extracellularly and the FLImP results reflect the extracellular geometry of ligand-bound EGFR complexes. A FLImP measurement of the separation of a pair of fluorophore-conjugated ligands bound to an EGFR complex (which is a result of the analysis of two-fluorophore spots based on criteria described previously[70,71]) produces an empirical posterior probability distribution of the separation[70]. As the separation cannot be negative, instead of a Gaussian, the empirical posterior takes the form of an asymmetric Rice distribution[266]. The width of the 68% confidence interval of the posterior distribution reflects the

precision of each pairwise separation measurement, which ultimately depends on the signal-to-noise ratio (***Figure 6-11***). The posteriors with 68% confidence intervals smaller than the required resolution (typically 4 –7 nm) are retained and pooled into a histogram (hereafter referred to as a FLImP distribution), from which one can derive structural information of EGFR complexes bound to more than one labelled EGFR ligands. From a FLImP distribution one can estimate the proportion of FLImP measurements consistent with particular species of EGFR complexes bound to more than one EGFR ligand as an indicator of the relative population of the species (will be shown later).

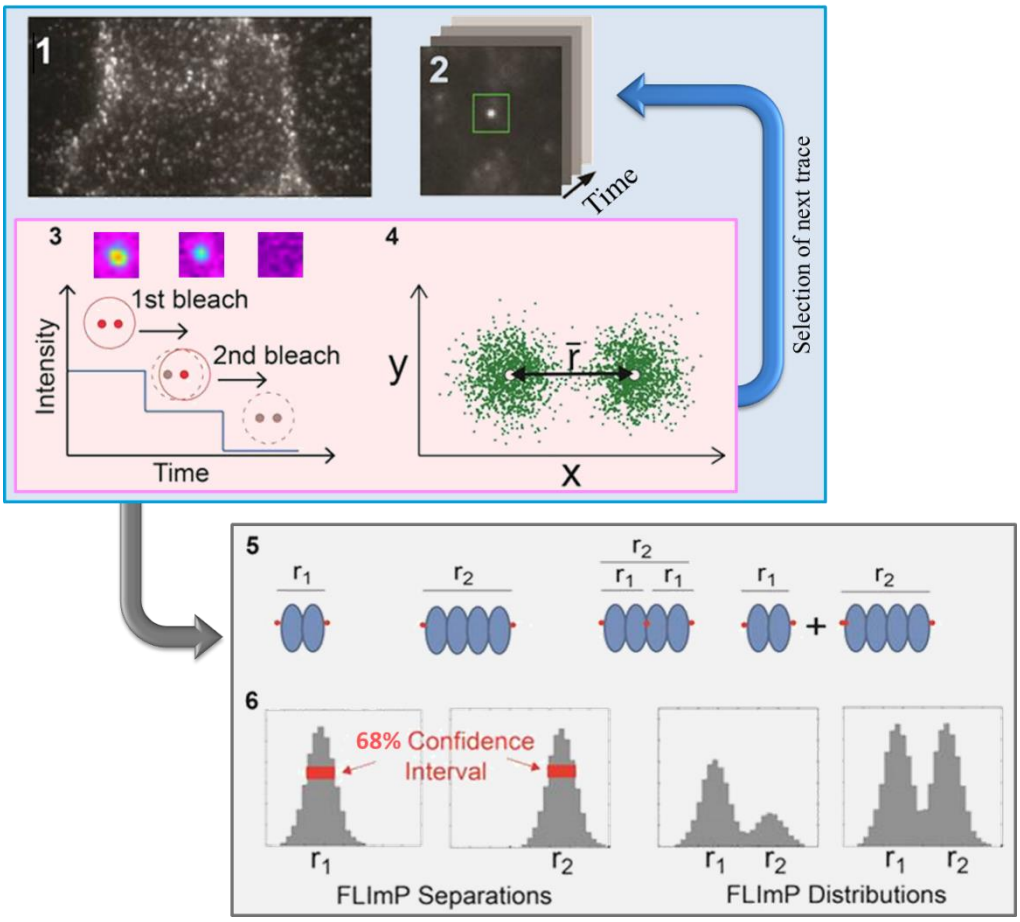


Figure 6-11: FLImP workflow. Steps to determine ligand separations using FLImP: (1) TIRF images are collected from fixed cells; (2) Spots corresponding to individual complexes are tracked to derive intensity *vs* time traces[96]; (3) A spot image of a molecular complex containing two fluorophores (each represented by a red dot) features two intensity levels and a decay to zero in two bleaching steps; when one fluorophore bleaches, the centroid position of the spot shifts; If more than two steps are present, the lowest two are analyzed (4) A global least-squares 7-parameter fit is used to identify the best intensity, x and y positions for each fluorophore and the full width at half-maximum of the point spread function, from which the separation between the molecules can be calculated with a precision determined by the localization error; (5) Example systems of a two-ligand EGFR dimer, a two-ligand tetramer, a three-ligand tetramer, and a mixture of a dimer and a tetramer; (6) The FLImP measurement of pairwise ligand separations obtained for each example system using FLImP, where the 68% confidence intervals are highlighted.

The size of the confidence interval depends on the combined localization errors of the two molecules. The FLImP measurements with confidence intervals smaller than the required resolution are pooled into a histogram, generating a so-called FLImP distribution that is well fitted by the sum of a discrete number of Rician peaks.

In order to be able to detect and characterize the presence of possibly two different (tethered and extended) dimeric populations with FLImP, the knowledge of the distances between fluorophore-conjugated ligands bound to both conformations is required (assuming that these distances are different). The lateral dimension of the unliganded extended ECD dimer is expected to be ~11nm, given the shape similarity of the unliganded ECD dimer seen in long-timescale MD simulations[63] and in EM images[252] with the active (EGF-bound) extended EGFR ECD dimer, for which the lateral dimension is ~11nm[34,35]. The ligand I have used for all FLImP experiment herein was the anti-EGFR Affibody which binds to the EGF epitope, as I discussed earlier (***Section 6.2***). Given that the estimated separation (based on crystallographic data) between the two fluorescent labels each bound to the two EGF in the back-to-back EGFR dimers is 12.5 nm[251], one would anticipate a similar distance between fluorophores-conjugated affibodies bound to extended dimers. Furthermore, L680N could also be used as a control for the FLImP measurements of the unliganded extended dimer since DOCA and SPT results suggested the presence of extended non-monomer complexes. While the conformation of the unliganded extended ECD dimer has been observed in both long-timescale MD simulations[63] and in EM images[252], the conformation of a tethered dimer is unknown. Furthermore, an estimation of the distance between fluorophores conjugated to ligands (with known epitope such as EGFR Affibody) cannot be made as the dimer interface is also unknown (even though EM images suggested intermolecular interaction between ECD domain IVs of a EGFR receptor where the ECD was fused to a dimeric immunoglobulin Fc domain[252]). Therefore, I decided to measure the distribution of separations, using FLImP, between labelled ligands of receptors to which the tethered conformation was reinforced.

Stabilisation of the tethered conformation or prevention of the ECD from adopting the extended conformation (or both simultaneously) can be achieved by either treatment with reagents targeting the ECD or by mutation. Attempts using the latter approach have been previously reported by Burgess and colleagues[263]. The authors mutated two key residues located at the centre of the auto-inhibited, intramolecular, interaction of the

tethered conformation (between domain II and IV) to cysteines in order to "lock" the receptor via a disulphide bond formation. However, inconsistency between different methods for determining the expression levels of this mutant led to the conclusion that the receptor might be partially misfolded[263]. On the other hand, treatment of EGFR with reagents such as antibodies (cetuximab[267] and matuzumab[24]) or nanobodies (7D12, EgA1 and 9G8)[162,268] has been successfully shown to alter the conformation of the receptor. Due to use of ECD probes that bind to either DI (EgB4) or DIII (Affibody and EGF) in the imaging experiments, I required a reagent which primarily promotes the tethered conformation but does not interfere with the binding of the ECD. Although, matuzumab, EgA1 and 9G8 meet both the criteria (matuzumab binds to domain II[24] whereas both EgA1 and 9G8 bind to the junction of domain II and III[268]), I selected the 9G8 nanobody (kind gift from Dr Paul van Bergen en Henegouwen, Utrecht University) for my experiments. Apart from the fact that nanobodies, in general, are smaller in size than antibodies (and hence produce less interference with other bound molecules), both EgA1 and 9G8 have been shown not only to prevent the formation of the extended conformation (as matuzumab does) but also to stabilise the tethered conformation[268]. Furthermore, 9G8 (***Figure 6-12A***) is 1.6-fold more potent than EgA1 as revealed by binding studies[268].

As mentioned above, 9G8 binding is not expected to interfere with the binding of the ECD probes (EgB4 and Affibody) as it binds to an epitope distinct from both EGF (Affibody is an EGF-competitive ligand[256,257]) and EgB4 (binds to DI[220]) epitopes. Nevertheless, I sought to characterize the binding of ECD probes by confocal microscopy in the absence and in the presence of a saturating concentration (200 nM) of 9G8 in order to ensure that their binding is not affected by 9G8 binding. For that, wt EGFR expressing cells were treated with or without 9G8 during labelling (1 h at 4^{o}C) with either EgB4-Alexa 488 or Affibody-Alexa 488 and then fixed. The binding of both probes to EGFR was assessed by comparing the frequency distributions of membrane pixel intensities (obtained by confocal microscopy) of the 9G8-treated with that of the untreated cells. As shown in ***Figure 6-12*** (***C*** and ***D***), the binding of both probes was unchanged, indicating that 9G8 does not compete with either probe and hence is a suitable reagent for my investigation.

In order to validate that 9G8 treatment generates higher number of tethered receptors than wt EGFR (and of course L680N EGFR) in cells I measured the distance of both DI

and DIII from the membrane. If more tethered receptors are formed, then the distance between ECD DIII and the membrane is expected to be higher than the distance of DI and the membrane. As shown in ***Figure 6-12B***, the distance from DI to the membrane in 9G8-bound receptors is 4.6 ± 0.8 nm, which is remarkably consistent with the 4.5 nm N-terminus-membrane separation predicted for the tethered conformation by MD simulations[255]. The distance from DIII, as expected, is longer (8.4 ± 0.6 nm) indicating the predominance of tethered receptors in the samples.

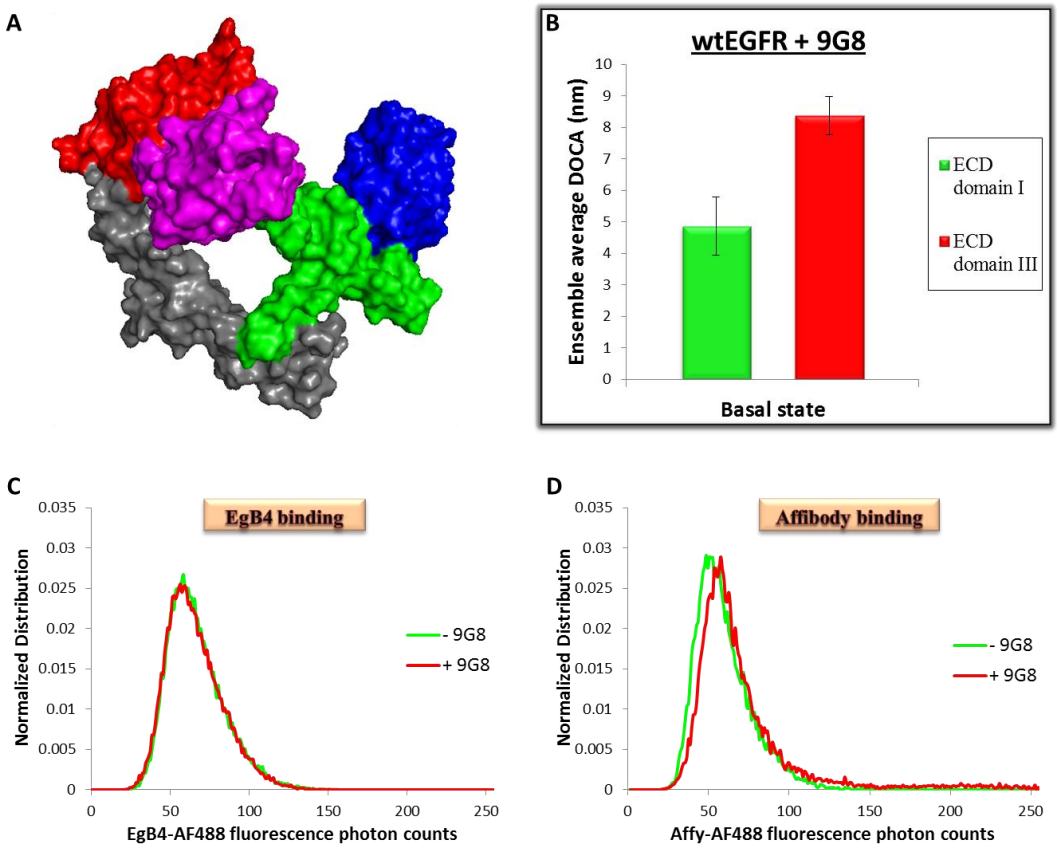


Figure 6-12: 9G8 nanobody binding promotes the tethered conformation. (**A**) Crystal structures of 9G8 nanobody bound to the ECD of EGFR in the tethered conformation (PDB: 4KRP). Colours: blue for sub-domain I, green for sub-domain II, red for sub-domain III, grey for sub-domain IV and magenta for 9G8. (**B**) Determination of DI and DIII distances (DOCA) in the basal state of wtGEFR pre-treated with 200 nM of 9G8 nanobody. (**C**) The binding of 50 nM EgB4-Alexa 488 to wt EGFR in the absence (green) or in the presence (red) of a saturating concentration (200 nM) of the 9G8 nanobody, assessed by the distribution of fluorescence photon counts per pixel of membrane (obtained with confocal microscopy). (**D**) Assessment of Affibody-Alexa 488 (20 nM) binding as in (**C**). All experiments were performed in triplicate.

As mentioned earlier, comparison of the fTKI and DOCA results between wt and the L680N EGFR led me to consider the likelihood of tethered dimers being present in the former receptor, in agreement with the findings of Bessman *et al.*[252]. Having

demonstrated that 9G8-treated receptors are predominantly tethered, I then employed the FLImP method in an attempt to characterize such tethered dimeric receptors. As a control for the extended conformation I used L680N EGFR which was shown earlier to form higher number of non-monomer complexes. For the FLImP experiments, cells were labelled with Affy-CF640R, similarly to DOCA, and then fixed. 9G8 treatment was performed simultaneously with the labelling.

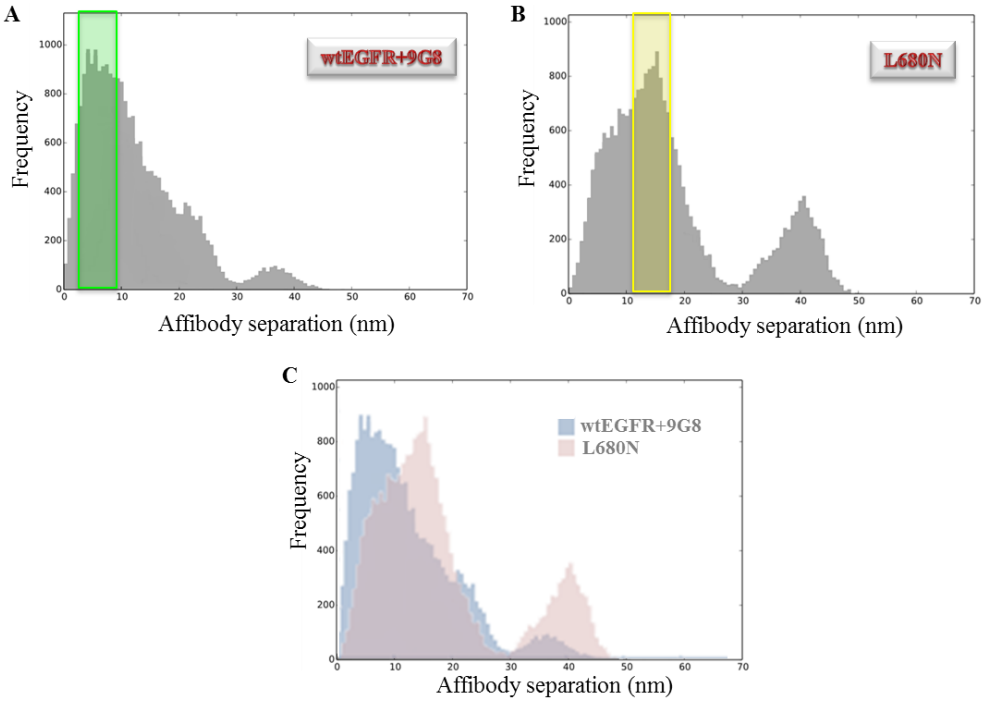


Figure 6-13: Determination of the separations of tethered and extended dimers. FLImP distributions (grey) of EGFR labelled with CF640R fluorophore-conjugated Affibody on the surface of CHO cells expressing either wt EGFR pre-treated with 200 nM of 9G8 (**A**) or L680N EGFR (**B**) labelled with 4 nM Affibody at 4°C, followed by chemical fixation. The FLImP distribution was compiled from ~30 FLImP measurements with confidence intervals <6 nm. (**C**) Merged FLImP distribution plots of (**A**) and (**B**). All experiments were performed in triplicate.

The FLImP distributions obtained from these conditions are shown in ***Figure 6-13***. Despite their overlap at the dimer range (***Figure 6-13C***), the shape of their distribution revealed two distinct peaks; whereas 9G8-treated receptors (***Figure 6-13A***) exhibited the highest densities at separations <10 nm (green box), L680N EGFR expressing cells displayed separations which more frequently appeared in the range of 10-18 nm (yellow box, ***Figure 6-13B***). Two separate peaks within these ranges were also recently reported[251] to appear, simultaneously, in the FLImP distribution of pairwise separations of basal wt EGFR (***Figure 6-14***, right panel). The authors further applied a Bayesian information criterion in order to determine the decomposition of the FLImP distribution

of Affibody separations into 4 (best fit) Rician peak components (***Figure 6-14***, left panel, highlighted by a green box). For the two main peaks, the centre was calculated to be at 4.1 (-1.9 +4.2) nm and at 12.0 (±1.6) nm (the smaller peaks at longer distances possibly reflect the presence of complexes larger than dimers). I previously speculated that the 4.1 nm separation was associated to the tethered conformation and that the 12 nm separation may reflect the distance between two Affibody molecules bound to an extracellular configuration akin to the back-to-back dimers. Collectively with the control experiments for the tethered (9G8-treated cells) and extended (L680N EGFR) dimers separations, the data suggest that these peaks represent those two dimeric populations.

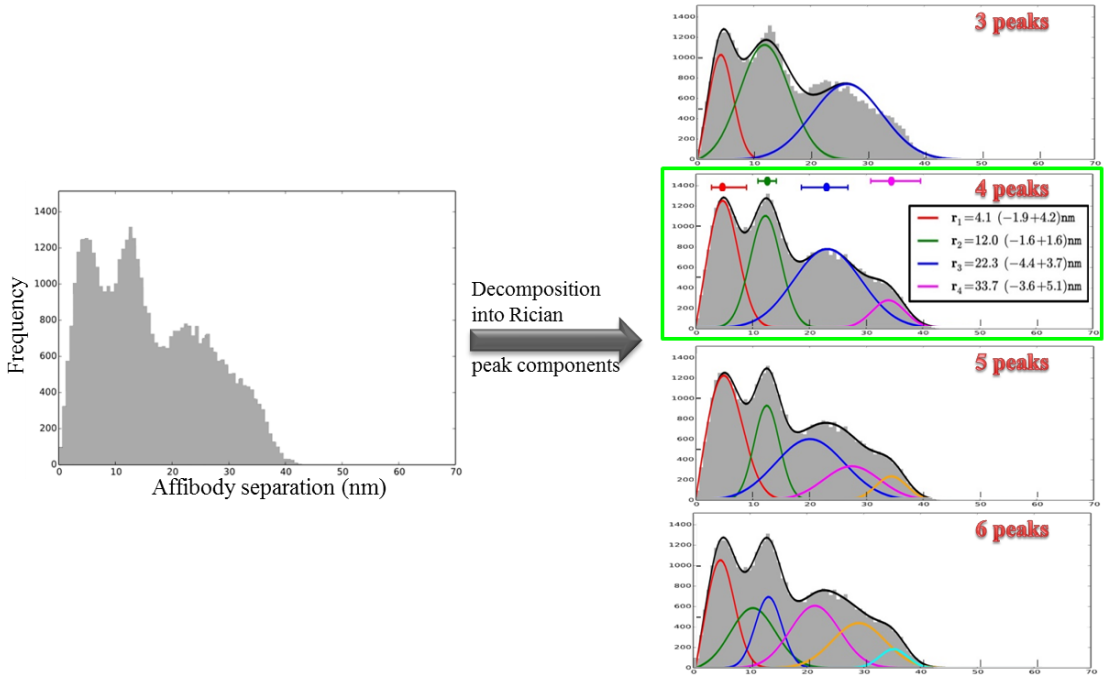


Figure 6-14: FLImP decomposition into its underlying Rician-peak components (adapted from Needham et al. Nat. Commun. 2016[251]**).** (Left panel) FLImP distribution (grey) of pairwise Affibody separations on CHO cells expressing wt EGFR treated with 4 nM Affibody-CF640R. (Right panel) Different fits using increasing numbers of peak components. The figure shows that increasing or decreasing the number of peaks from 4 (green box) is not objectively justified by the data. The distribution is fitted (black line) with a sum of 3-6 Rician peaks (colour lines). The number of peaks used was determined using a Bayesian information criterion. The best-fit positions of the peaks and error bars are shown in the inset. The errors in the fit were calculated as described in Needham *et al.*[251].

In order to be able to correlate the effect of each ECD dimeric population with the TKD conformation (and hence receptor activation), their relative proportion in each mutant or condition was calculated. To do this, I calculated the percentage of

distributions consistent with FLImP measurements in the ranges of either 4.1 (-1.9 +4.2) nm (green coloured area) or 12.0 (±1.6) nm (yellow coloured area), based on the ranges derived from the peaks of the basal wt EGFR (***Figure 6-15***). Distributions outside the dimer range (>13.6 nm) were also calculated (shown in blue). Hereafter the ranges at 4.1 (-1.9 +4.2) nm and at 12.0 (±1.6) nm are referred to as tethered and extended dimers, respectively.

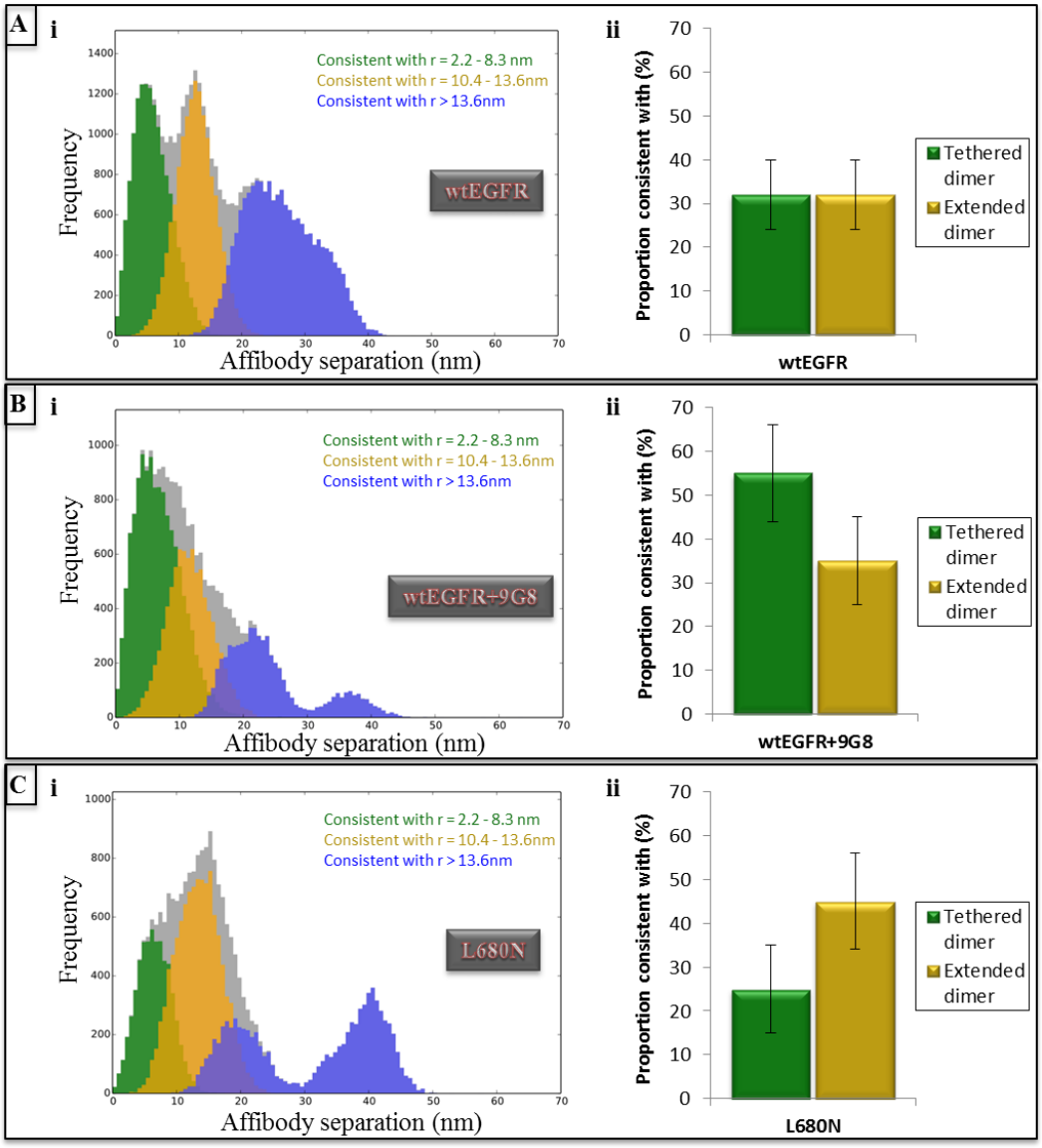


Figure 6-15: Relative populations of the tethered and extended dimers of basal wt EGFR, 9G8-treated wt EGFR and L680N EGFR expressing cells. FLimP results obtained from CHO cells expressing wt EGFR (**A**) or wt EGFR pre-treated with 9G8 (**B**) or L680N EGFR (**C**), all labelled with Affy-CF640R. Plots denoted as (i) show the FLImP distribution (grey) and the distributions (green, yellow or blue) compiled from the FLImP measurements whose ranges of 68% confidence overlap with the ranges of Affibody separations of the expected tethered (4.1 (-1.9 +4.2) nm, green) or extended (12.0 ± 1.6 nm, yellow) dimer or higher order oligomers (>13.6 nm, blue). Relative populations of the tethered and extended dimers determined from the FLImP measurements shown in (i) are illustrated in (ii). For each construct, the tethered percentage is estimated by the ratio of the green integral area to the

integral area of all FLImP measurements whose 68% confidence intervals overlaps with the three regions (green, yellow and blue). Likewise the extended dimer population is calculated based on the yellow area. Oligomer populations are calculated in the same way, using the blue instead of the green integral area. Error bars were calculated by bootstrap-resampling the data 1000 times with replacement and repeating the analysis.

As shown in ***Figure 6-15A***, in the basal unliganded wt EGFR both the tethered and extended ECD dimers coexist in similar levels on the cell surface. Stabilization of the tethered conformation (9G8 binding) or disruption of the asymmetric dimer interface (L680N EGFR) significantly alters the relative populations: whereas 9G8 binding to unliganded wt EGFR enhances the number of separations consistent with the tethered dimer (***Figure 6-15B***), L680N point mutation favours the formation of extended dimers (***Figure 6-15C***).

The presence of higher order oligomers in the wt receptor was also observed (***Figure 6-15***, blue). The alterations from the wt receptor affected both the overall percentage of the oligomers (***Figure 6-16***) and the frequency of separations at these ranges. Oligomers detected in basal wt EGFR appear more frequently at separations which could possibly be explained by tetramer (20-30 nm) or hexamer (30-40 nm) formation. By comparing 9G8 treatment and L680N EGFR the data suggest that Affibody separations consistent with the former range appears to be asymmetric dimer-dependent whereas promotion of extended ECDs and symmetric dimers (L680N mutant) favours the formation of even larger complexes (peak at ~40nm), consistent with SPT results.

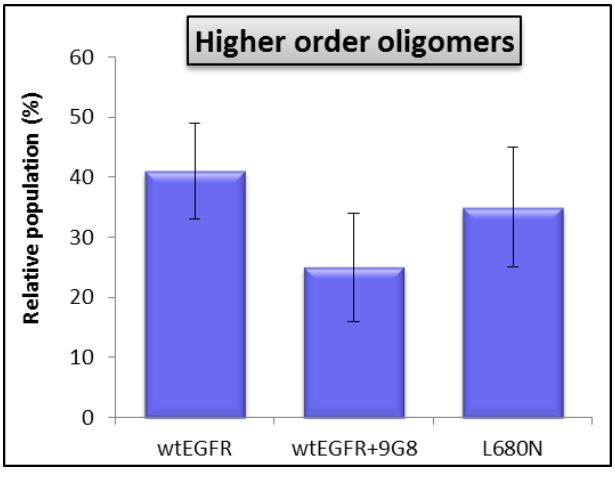


Figure 6-16: Relative populations of the higher order oligomers of basal wt EGFR, 9G8-treated wt EGFR and L680N EGFR expressing cells. Relative populations of the higher order oligomers determined from the FLImP measurements shown in ***Figure 6-15*** (**A**, **B** and **C**, left panel plots (i)) which were calculated as earlier.

6.4 Outside-in and Inside-out Conformation Coupling of Tethered ECD and Asymmetric ICD in the Basal Receptor

In the previous section I showed that 9G8 treatment not only lock the receptor in the tethered conformation, in agreement with previously reported data[268], but also favours the formation of tethered dimers which have been also recently visualized by EM[252]. Furthermore I showed that these tethered dimers are destabilized when the asymmetric dimer interface is disrupted (with concomitant promotion of symmetric dimers). All these suggest a linkage between the ECD tethered unliganded dimers with the asymmetric TKD dimers present in the basal receptor. In order to investigate this conformational coupling in the basal receptor state I used two approaches:

- Outside-in receptor regulation by promoting the ECD tethered conformation
- Inside-out signalling by promoting asymmetric dimers

6.4.1 Forcing the Tethered ECD Conformation

Even though 9G8 treatment was shown by FLImP to favour the formation of tethered ECD dimers, the overall relative population of higher order oligomers was significantly lower than in untreated cells (25±9% and 41±8% respectively, ***Figure 6-16***). This observation could suggest either that promotion of the tethered configuration favours dimeric complexes (or disrupts oligomerization) or that the monomer/non-monomer ratio is higher than in basal wt EGFR. If the latter is true, then the number of tethered dimers will not be notably different between the basal and the 9G8-treated receptors and hence the investigation of the TKD conformation might prove difficult. For this reason I employed the SPT method described earlier for determining the correlation between the tethered conformation and the formation of non-monomer complexes.

The experiment was performed in the same manner as that described in ***Section 6.2.1***, but this time in the presence of a saturating concentration (200 nM) of 9G8 nanobody. The results are shown in ***Figure 6-17*** (basal wt EGFR is also shown for comparison). SPT experiments revealed that 9G8 induced a small but significant ($p < 0.05$) increase (~1.6-fold) of non-monomer complexes compared to untreated wt EGFR

cells (***Figure 6-17A***). This is especially important considering that FLImP results suggest that the non-monomer fraction is mainly populated by dimers. Furthermore, the duration (T_{on}) of these interactions was also increased (***Figure 6-17B***). These results demonstrate that tethered ECDs are able to interact via well-defined extracellular interfaces capable of driving dimerization. It is also possible that these dimers gained additional stability by the formation of asymmetric kinase dimers, if the tethered ECD dimers are indeed coupled to this TKD configuration.

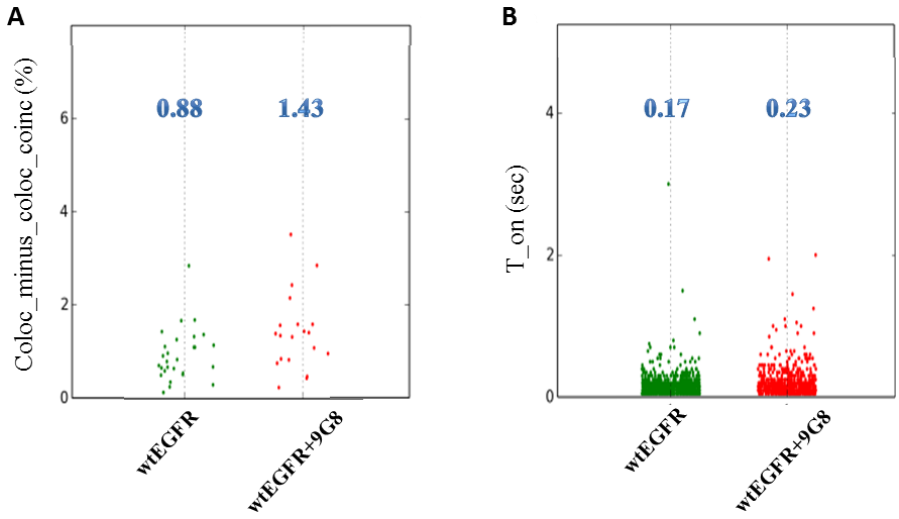


Figure 6-17: Comparison of the fraction of colocalized tracks (A) and duration of the colocalization events (B) between basal and 9G8-treated wt EGFR cells. Cells expressing wt EGFR were labelled with Affibody molecules conjugated to Alexa-488 and CF640R in the absence (green) or presence (red) of 9G8 nanobody. All experiments were performed in triplicate.

Having showed that 9G8-bound receptors form stable non-monomer complexes I then sought to determine the conformation of their TKDs. For that I employed the fTKI method. As discussed and demonstrated earlier (***Figure 5-12***), f-Gef binding is increased proportionally with the number of active ATP-clefts (which can only be formed after allosteric activation). On the other hand, type II fTKIs (Lapatinib and Neratinib) binding is inversely proportional to activation (***Figure 5-16***). Lastly, f-Pelitinib binding is conformationally-independent as it was demonstrated to bind both inactive and active kinases with the same affinity (***Figure 5-14***). Therefore, if tethered ECD dimers are conformationally coupled to asymmetric kinase dimers then 9G8-treated cells will exhibit higher membrane signal when labelled with f-Gefitinib and lower when type II fTKIs are used, compared to untreated cells. Cells for these experiments were prepared as previously described for the basal wt receptor, (***Section***

5.3.1) but in the presence of 200 nM of 9G8 during the ECD labelling step. The results are shown below (***Figure 6-18***)

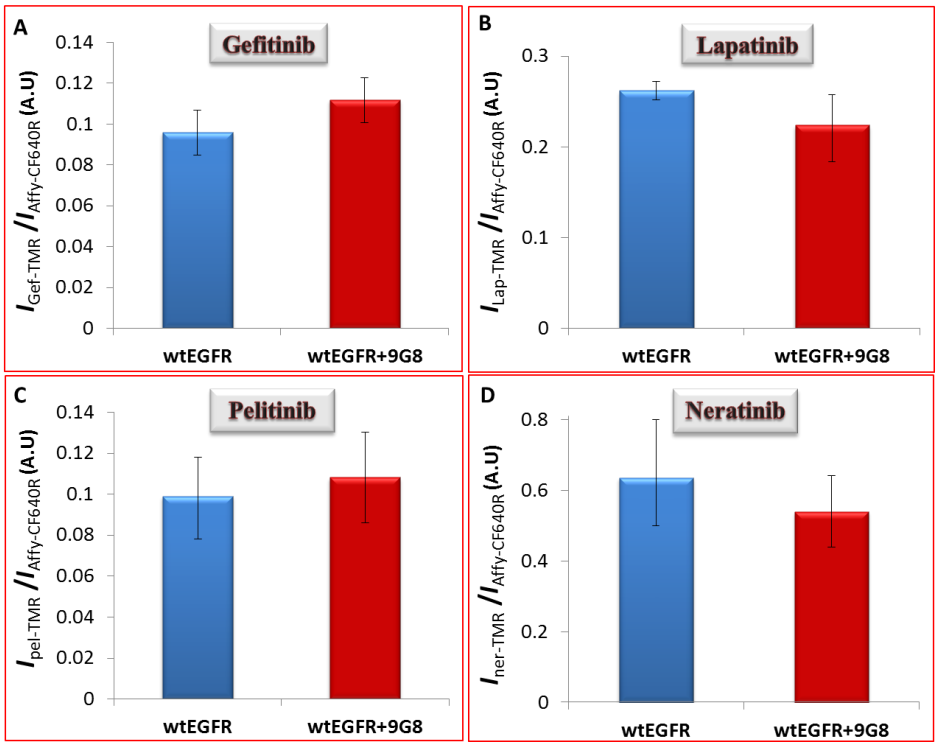


Figure 6-18: fTKIs binding in CHO cells expressing wt EGFR pre-treated with 200 nM of 9G8 nanobody and comparison with untreated cells. fTKIs binding in wt EGFR-expressing CHO cells treated with saturated concentration (200 nM) of 9G8 nanobody (red column bars) and comparison with their binding in untreated cells (blue column bars). Affy-CF640R was used to label the ECD in both experiments. Gef-TMR ($P > 0.05$) data is shown in (**A**), Lap-TMR ($P > 0.05$) in (**B**), Pel-TMR ($P > 0.05$) in (**C**) and Ner-TMR ($P > 0.05$) in (**D**). Data were analyzed as discussed in ***Section 5.3.1***. All experiments were performed in triplicate.

The results from fTKIs binding experiments appear to be consistent with my hypothesis: 9G8 treatment led to an increase of f-Gef membrane signal and a decrease of both f-Lap and f-Ner binding. Furthermore, the f-Pel binding appears to be unaffected. As the changes after 9G8 treatment were relatively small, an unpaired, two sample t-test was used to objectively compare the mean fTKI binding intensity ratio of the four fTKIs before and after 9G8 treatment (performed using OriginPro). The determined P values showed that there was no difference between any of the untreated and treated means significant to the 0.05 level (0.29 for f-Gef, 0.26 for f-Lap, 0.78 for f-Pel and 0.13 for f-Ner) so I consider the results of these experiments to be inconclusive.

An alternative method for determining whether the number of active TKDs has increased as a result of 9G8 treatment is to determine receptor phosphorylation. Western

blots analysis of 9G8-treated CHO cells expressing wt EGFR from unpublished data (Zanetti-Domingues *et al.*, manuscript in preparation) showed that phosphorylation was 1.4-fold higher than that of the basal (untreated) receptor. The reason why the observed increased in receptor phosphorylation was not able to be detected by the fTKIs could be due to the fact that in an asymmetric TKD dimer unit one TKD adopts the inactive and one the active conformation (***Figure 6-19***).

FLImP analysis of 9G8-treated cells showed that promotion of unliganded tethered receptors favours the formation of dimers, probably at the expense of higher order oligomers when it is compared to untreated wt EGFR expressing cells (***Figure 6-15B***). Crystallographic data have revealed that unperturbed homodimeric TKDs of both EGFR[49] and ErbB4[51] form a daisy chain of asymmetric dimers where the receiver kinase of one dimer is also the activator kinase of the following dimer (***Figure 6-19A***). In this arrangement which occurs in EGF-stimulated receptors as recently shown[265], only the activator kinase on one end of the daisy chain adopts the inactive conformation (recognized by type II and irreversible type I TKIs) while the rest of the kinases (here shown in a tetrameric configuration) are active (recognized only by type I TKIs). In this case, type II fTKIs would bind only to ¼ of the population of EGF-stimulated receptors, if only tetramers existed as illustrated in ***Figure 6-19A***. On the other hand, TKDs which can form a dimeric asymmetric unit but not the daisy chain, as in the case of the asymmetric EGFR/ErbB3 heterodimer which cannot form TKD oligomers[269], half of the receptor population would adopt the inactive TKD conformation (***Figure 6-19B***). Even though such a dimeric unit would be phosphorylated, both type I and II fTKIs would be able to bind. In cell context, the number of active TKDs of such receptors which exhibit the aforementioned 1.4-fold increase in phosphorylation would only slightly change.

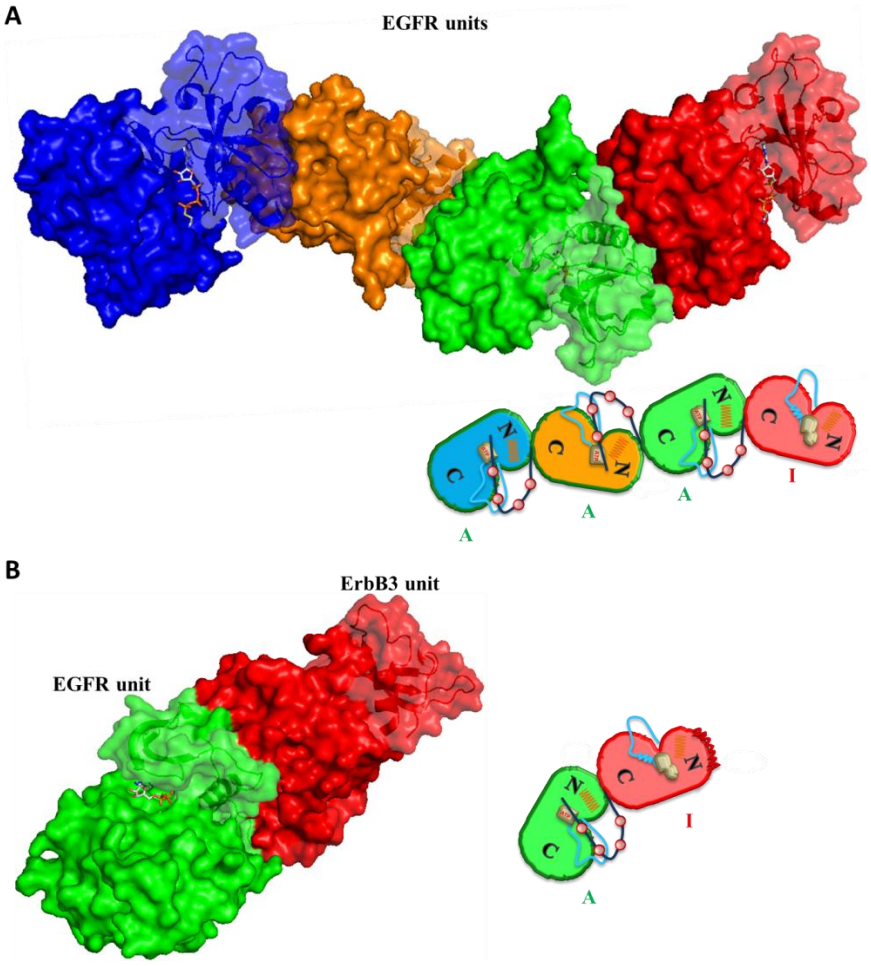


Figure 6-19: EGFR asymmetric homodimer daisy chain and EGFR-ErbB3 asymmetric heterodimer. A) In all the previously reported crystalized asymmetric kinase dimers (for both EGFR[49] and ErbB4[51] homodimer), the molecules form a daisy chain where the receiver of one asymmetric unit is also the activator of the following monomer (here is the PDB: 2GS6). In this configuration, all the ATP-cleft form the active conformation. **B**) In the recently resolved EGFR-ErbB3[269] crystal structure (PDB: 4RIW), the configuration is absent due to the presence of ErbB3 in the activator position.

In conclusion, the data are consistent with the assumption that tethered unliganded receptors are conformationally coupled with asymmetric dimers. Most importantly I have found that, contrary to expectation in the field, promotion of the tethered ECD conformation results in the increase of active kinases (and hence increased phosphorylation, Zanetti-Domingues *et al.* manuscript in preparation) in the absence of ligand. A literature search revealed a similar finding for a E578C EGFR mutant in which the tethered conformation was promoted by improving the packing of the ECD domain II/domain IV interface (intramolecular auto-inhibited interaction of the tethered conformation) and resulted in higher basal phosphorylation (which however was not commented upon the authors)[263]. "Locking" the ECD conformation to the tethered one

has been proposed as a mechanism for EGFR signalling inhibition due to the assumption that tethered receptors would be monomeric and hence inactive[268]. Nanobodies such as 9G8 and 7D12 and antibodies such as matuzumab have been developed based on that mechanism and in fact these molecules has been shown to successfully suppress EGFR phosphorylation after EGF stimulation. However, my data suggest that a population of ligand-free receptors might not be actually inhibited. The relevance of this finding to cell signalling is not understood. Further studies using such molecules and other cell lines need to be conducted to determine whether the effect constitutes a general pattern or not.

6.4.2 Promoting EGFR Asymmetric Dimers in the Basal State

In order to further investigate the presence of unliganded EGFR dimers where the ECDs adopt the tethered conformation coupled with the TKDs an asymmetric configuration, an alternative approach was employed. As opposed to promoting the ECD tethered conformation, I sought to increase the number of the unliganded dimers in cells by inducing asymmetric dimerization of the ICDs in the absence of ligand.

As briefly discussed in ***Section 4.2***, several studies using either cross linking (BS^3) methods[65,217,235,270,271] or FRET measurements[137,230,235] have shown that cells treated with type I TKIs (either reversible or irreversible) exhibited higher number of EGFR dimers compared to untreated cells. Furthermore, these receptors were found to form asymmetric dimers with their ECDs displaying a conformation more akin to that of the tethered monomer in the absence of ligand in EM studies performed on detergent-solubilized receptors[217]. Therefore I decided to reproduce these experiments in cells to further study the conformational coupling between the ECD and the TKD of the unliganded receptor seen in 9G8-treated cells (***Section 6.4.1***).

All the following experiments on inhibitor-induced dimerization were carried out in CHO cells expressing wt EGFR. In order for any inhibitor to induce the associated conformational change, receptors need to be treated in live cells with a saturating concentration of an unlabelled TKI so that most receptors exist in an inhibitor-bound state. Since TKIs in such a concentration inhibit EGFR phosphorylation, the formation of asymmetric dimers in cells cannot be verified by WB analysis as previously discussed for 9G8 (***Section 6.4.1***). Therefore the presence of asymmetric dimers can

only be investigated using the fTKI method. However this is a quantitative method which is based on the recognition of unoccupied ATP sites (either inactive or active) from the fluorescent inhibitors; hence the type I TKI used to induce dimerization needs to be reversible so it can be washed out after fixation. Even though any type I reversible TKI could in principle be used in the inhibitor-induced dimerization experiments, Gefitinib® was chosen as, apart from being the parent compound of my fluorescent type I reversible inhibitor, it was also the compound used in the EM study reported by Springer and colleagues[65,217]. Inhibitor treatment (1 μM) in the following experiments was performed during starvation (2h) as well as during ECD labelling (1 h for methods in fixed cells). For live-cell experiments (SPT), the same concentration of the inhibitor was kept in the media while acquiring data.

The results from the fTKI binding experiments for determining the TKD conformation of the Gefitinib-treated wt receptors are shown in ***Figure 6-20***. Even though, based on previous studies, one would expect that f-Gef membrane signal to be higher in inhibitor-treated cells (formation of asymmetric dimers) than in untreated basal receptors (***Figure 6-20A***, red and blue respectively), analysis of the data revealed a reduction in f-Gef binding by ⅓ compared to basal receptor. This reduction could be either due to a conformation different from the expected, in which more inactive TKDs have been generated as a result of Gefitinib treatment, or due to competition of f-Gef with residual unlabelled Gefitinib. If the former were true, the binding of both type II fTKIs would be increased. However, in both cases the binding was lower compared to untreated cells (***Figure 6-20 B&D***), a result which supports the latter hypothesis and demonstrates an increase in the active TKD population. Furthermore, the fact that f-Pel binding remained unaffected (***Figure 6-20C***) indicates that the ATP cleft is fully accessible and in a conformation which can be recognized at least by f-Gef (as another type I fTKI). Therefore, the reduction of f-Gef binding could be attributed to competition with unlabelled Gefitinib.

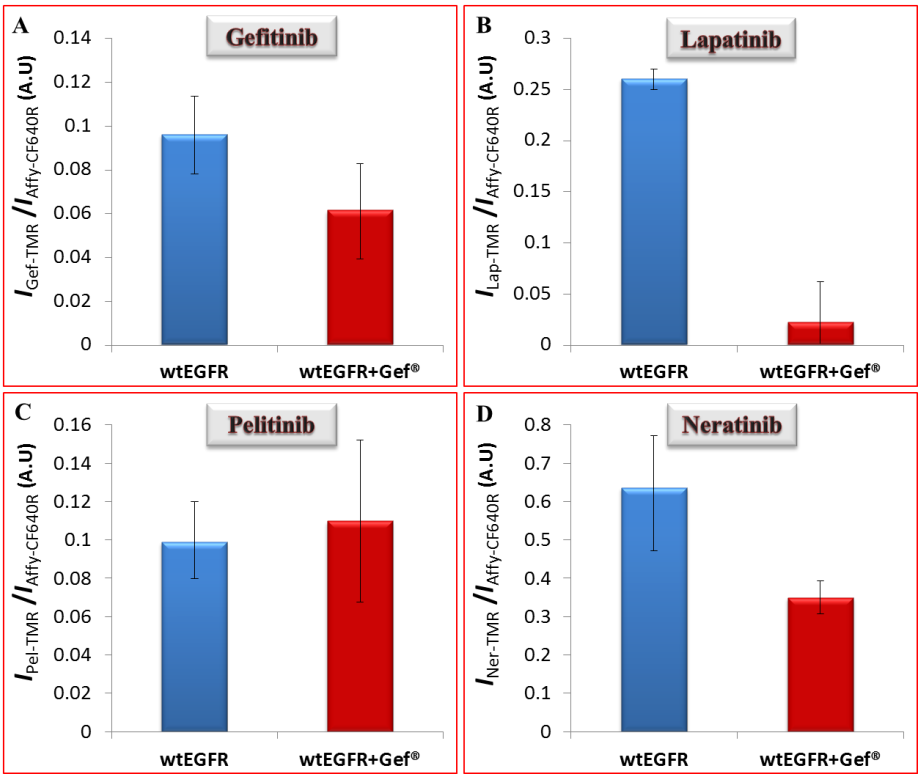


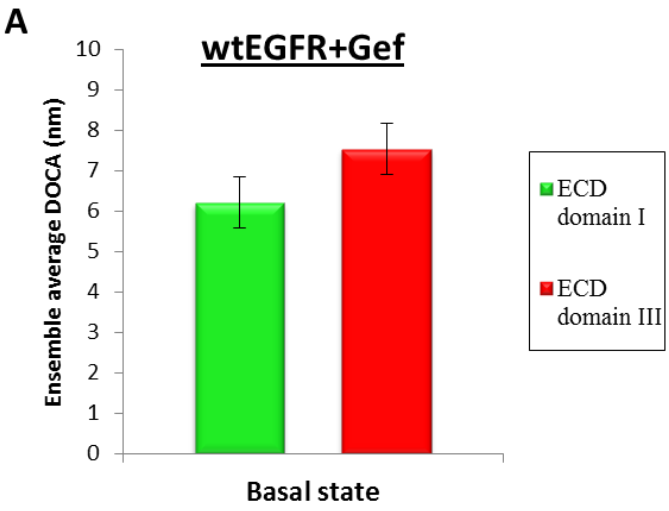
Figure 6-20: Determination of the TKD conformation of wt EGFR pre-treated with Gefitinib®. fTKIs binding in CHO cells expressing wt EGFR pre-treated with 1μM of Gefitinib® (red column bars) and comparison with their binding in untreated cells (blue column bars). Affy-CF640R was used to label the ECD in both experiments. Gef-TMR data is shown in (**A**), Lap-TMR in (**B**), Pel-TMR in (**C**) and Ner-TMR in (**D**). Data were analyzed as discussed in ***Section 5.3.1***. All experiments were performed in triplicate.

The fact that f-Pel binding was not reduced by residual unlabelled Gefitinib suggests that the concentration of the latter is minimal, which is in agreement with what was observed when live-cell labelling with reversible fTKIs was attempted (***Section 5.2***). While competition between reversible compounds (f-Gef and Gefitinib®) is a continuous process, displacement of a reversible compound by an irreversible one only occurs once. The same differences were also observed in type II fTKIs experiments; while f-Lap binding was almost abolished (combination of competition and reduced number of inactive kianses), f-Ner membrane signal was reduced by 50%. An explanation for the fact that f-Ner binding was not abolished as a result of the formation of asymmetric dimers may lie in the presence of mostly dimeric receptors with asymmetric TKD dimer configurations which have been promoted due to Gefitinib® treatment. As discussed in ***Section 6.4.1***, promotion of the tethered conformation favours dimerization. Furthermore, such dimers appear to be conformationally coupled to asymmetric dimers (outside-in regulation). If promotion of asymmetric dimers

(Gefitinib® treatment) is also accompanied by formation of tethered ECDs as previously seen in EM[217] (inside-out regulation), then Gefitinib-treated cells are expected to be highly dimeric (will be also investigated later). If that is the case then type II fTKIs would be able to bind to 50% of the total number of receptors which is remarkably consistent with the f-Ner result.

Having found that Gefitinib-treated receptors form higher number of asymmetric dimers, the next step was to determine the conformation of the ECD in the same conditions. For that, the three methods for probing the ECD structure used earlier, (DOCA, SPT and FLImP, were employed. The results are shown in ***Figure 6-21***.

As in the case of 9G8 experiments (***Figure 6-12***), DOCA measurements of Gefitinib-treated wt EGFR expressing cells suggest an increase in the number of tethered complexes compared to untreated wt EGFR (***Figure 6-21A***). However the distance of both DI and DIII, as well as the difference between them, were different from those obtained from 9G8-bound receptors (6.2 and 7.54 nm as opposed to 4.87 and 8.37 nm for 9G8 treatment). A possible explanation for these differences might be due to different mechanisms of formation, and hence stabilization, of the ECD conformation. 9G8 binds to the receptor and "locks" it in the tethered conformation whereas in Gefitinib-treated cells the promotion of the tethered conformation is driven from inside due to the conformational coupling of asymmetric dimers with the unliganded tethered ECD dimers. Another explanation might be due to fluctuations of the tethered ECD conformation as opposed to 9G8-bound receptors.



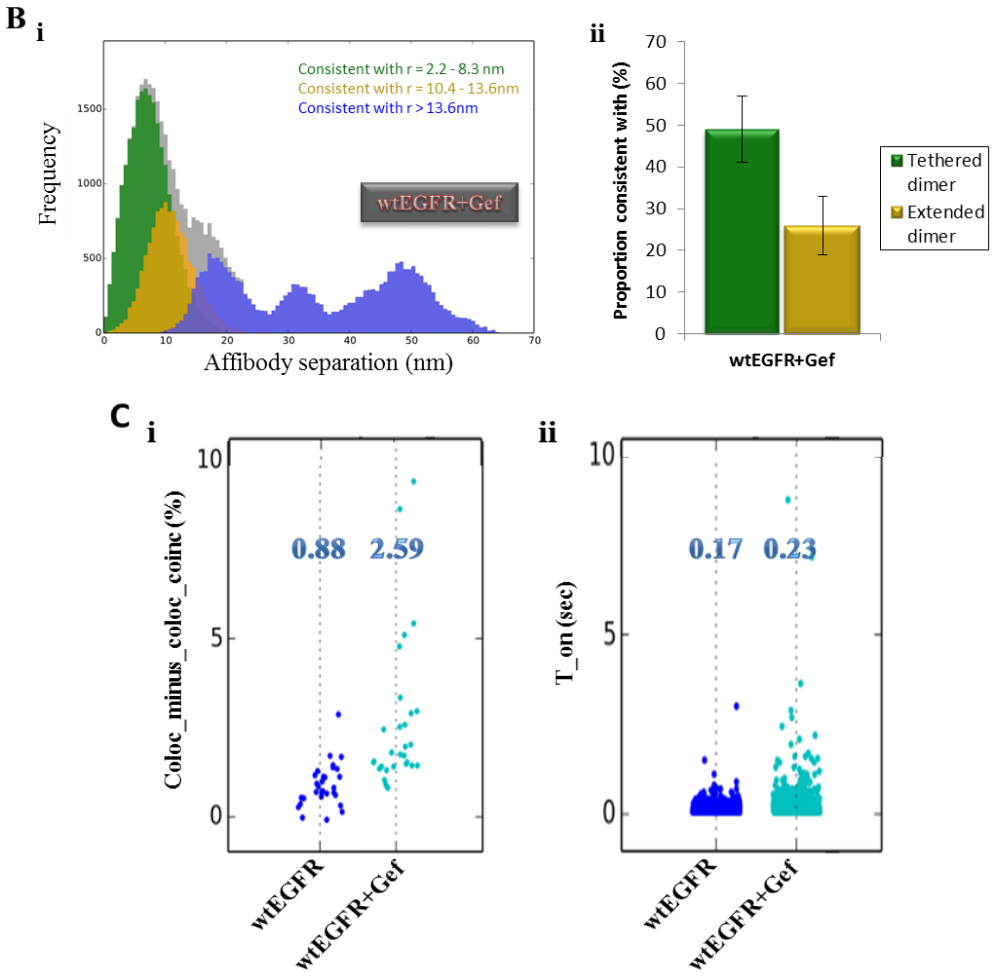


Figure 6-21: Determination of the ECD conformation and configuration of wt EGFR pre-treated with Gefitinib®. DOCA measurement (**A**), FLImP results (**B**) and SPT analysis (**C**) of wt EGFR expressing CHO cells pre-treated with 1 µM of Gefitinib®. All experiments were performed in triplicate.

The results from FLImP analysis were also consistent with an increase presence of tethered ECD dimers in Gefitinib-treated cells was also confirmed by FLImP analysis (***Figure 6-21B***). In the case of inhibitor-induced dimerization the ratio of the fraction of separations consistent with tethered/extended dimers (***Figure 6-21B, ii***) was increased compared to 9G8-treated cells (1.88 and 1.57 respectively). This additional stabilization of the tethered dimers when dimerization is driven by the formation of the well-defined asymmetric dimer interfaces (Gefitinib treatment) was also confirmed by 2-colour SPT colocalization analysis (***Figure 6-21C, i***). Promotion of asymmetric dimerization resulted in a 1.8-fold increase of the mean colocalized fraction of non-monomer complexes compared to promotion of tethered conformation (9G8). Despite the increased colocalization frequency, the duration of the interactions (T_{on}, ***Figure 6-21C,***

ii) calculated from both treatments was identical which might reflect the absence of stable oligomers.

FLImP distribution and SPT colocalization analysis also suggest that Gefitinib, as 9G8, favours the formation of dimers within the overall non-monomeric population. This observation is in agreement with the previous assumption that receptors treated with Gefitinib (and likely other type I TKIs) form asymmetric dimers but, in contrast with EGF-stimulated receptors[265], oligomerization is disrupted, despite the fact that TKIs are not expected to affect the interfaces required for oligomerization. The reason for this disruption, as the present data suggest, might be due to inability of the receptors to form stable higher order oligomers when the ECD adopts the tethered conformation.

In conclusion, in this section the previously proposed inside-out regulation of EGFR[65,103,217] was exploited to further investigate the conformation coupling of the "active" population of the basal receptor seen in 9G8-treated cells. It was found using the fTKI method that Gefitinib, a type I TKI, promotes asymmetric dimerization in cells consistent with previous results for the solubilized receptor[217]. Furthermore, it was shown that this TKD conformation in unliganded receptors promotes the formation of tethered ECD dimers, consistent with the data shown in ***Section 6.4.1***. Formation of asymmetric kinase dimers appears to be more crucial in stabilizing dimers with a tethered ECD configuration than the formation of the tethered conformation itself, or the destabilization of the extended conformation, since monomers are believed to be tethered[36,65]. A scheme summarizing the findings from ***Chapter 6*** regarding the architecture of the basal EGFR in cells is shown below (***Figure 6-22***).

Basal wt EGFR exists in equilibrium between three different conformations; inactive monomers, inactive extended dimers and active tethered dimers. The two dimeric populations were shown to coexist in similar levels when the receptor is unperturbed. The extended dimers appear to be associated with higher order oligomers, in agreement with what was recently reported by Needham *et al.*[251] regarding EGFR oligomerization due to interactions between extended ECDs, whereas the tethered dimers disfavour higher order oligomers. This may constitute an auto-inhibition mechanism to prevent aberrant activation of the receptor in the basal state. Further studies are required to determine whether the two dimeric populations are linked and, most importantly, which of the three conformations is important in the transition to the active receptor upon ligand binding.

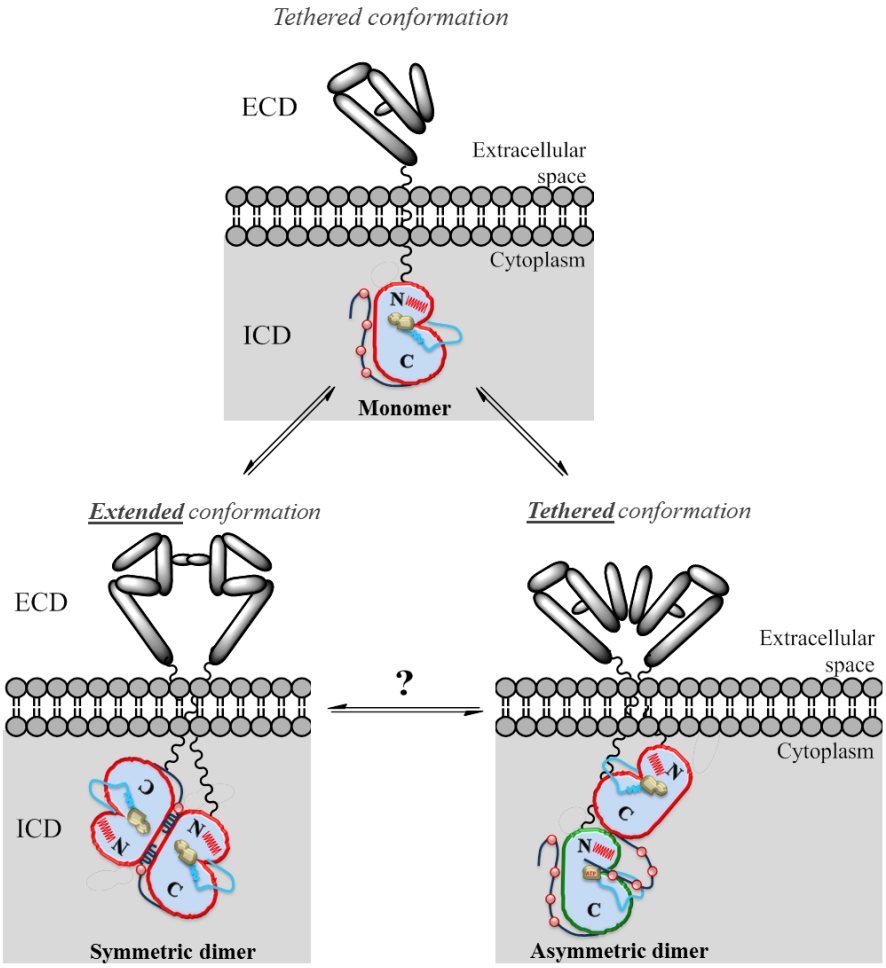


Figure 6-22: Proposed architecture of EGFR on the cell surface in the basal state.

7 Insight into Activation of Non-Small Cell Lung Cancer EGFR Mutants and Drug Resistance

7.1 Introduction

Physiological regulation of ErbB family members is often disrupted in various cancer types, leading to hyperactivation or ligand-independent activation. This deregulation could be the result of activating mutations, amplifications or overexpression (usually EGFR or ErbB2).

Activating mutations on all ErbB members have been found in all five structural elements participating in the asymmetric dimerization: ECD[272,273], TM[80], JM[53], TKD[274] and C-terminal tail[275]. Unfortunately, the structural basis of the activation of some of these mutations which could potentially lead to the design of more effective therapeutics remains unclear. Furthermore, mutations on both the ECD[276] and TKD[223] have been shown to confer resistance to different TKIs. Herein, I will investigate the role in EGFR activation and drug resistance of the two most common mutations occurring in non-small-cell lung cancer (NSCLC): the primary exon 21 point mutation (L858R or L834R) and the secondary exon 20 point mutation (L858R/T790M),.

7.2 Effect of L858R and L858R/T790M on EGFR Architecture

NSCLC constitute the most frequently type of lung cancer[222] which is often driven by mutations on the TKD of EGFR. These somatic mutations are caused by in-frame deletion (Δ747-749 and Δ746-750)[277], point mutation (either single mutant such as L858R[25], G719S[25] and T790M[213] or double mutant such as L858R/T790M[223] and G719S/T790M[278]) or in-frame insertion (D770_N771insNPG[279]).

Leu[858] is located at the N-terminal region of the A-loop, adjacent to the small helical DFG motif, and in the inactive state of the TKD participates in a set of hydrophobic interactions which hold the αC-helix outwards. Crystal structures revealed that mutation to arginine results in disruption of these interactions which in turn allows the αC-helix

to adopt the position that favours the active conformation[25] (equivalent mutation has been also reported for ErbB3[27]). Furthermore, molecular dynamic simulations suggested that the mutation also promotes dimerization by reducing the local disorder of the N-lobe[280]. Thr[790] is located in the hinge region and is also known as the "gatekeeper" residue. Its mutation also involves displacement of the DFG motif and stabilization of the active conformation. The double L858R/T790M has been suggested to further stabilize the active conformation by reducing the energy barrier for the transition from the inactive to the active conformation as well as by stabilizing the αC-helix, and subsequently the TKD, in the active conformation[281].

Despite these findings, the effect of these TKD mutations on the ECD structure, for example through inside-out regulation, and on the geometry of the receptor on cells was only recently explored[103]. Valley *et al.*[103] showed that L858R adopts an extended conformation in the absence of ligand which was also accompanied by increased dimerization and aggregation. In order to get a better insight into the mechanism that drives the L858R and the L858R/T790M mutations I decided to characterize the NSCLC mutants using the methods described in ***Chapter 6***. For these experiments wt CHO cells were transiently transfected (optimization of the transfection was also performed) with the desired plasmid using ViaFect (Promega®) and a reverse transfection as described for ErbB2-4 in ***Section 5.2***. Their expression levels were found to be identical as assessed by confocal microscopy.

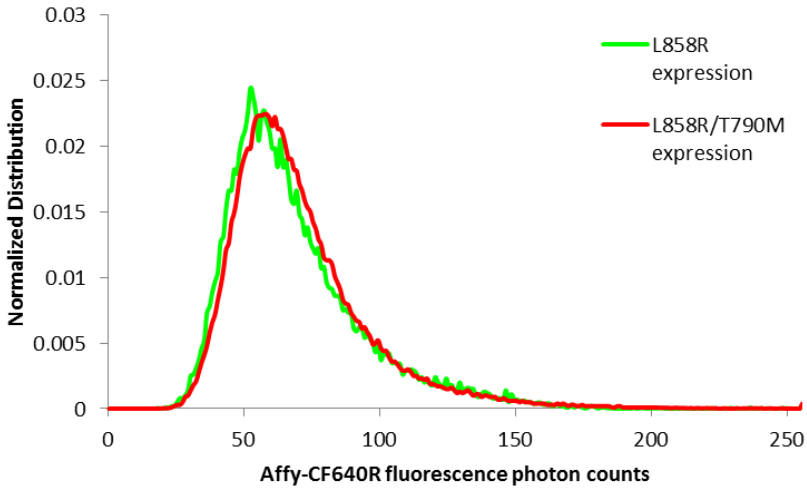


Figure 7-1: Assessment of expression levels of L858R and L858R/T790M in wt CHO cells. wt CHO cells expressing either the L858R (green line) or the L858R/T790M EGFR mutant were labelled with a saturating concentration (100 nM) of Affy-CF640R and then fixed. Assessment of the expression of both NSCLC mutants was performed by confocal microscopy as described in ***Figure 6-12C*** & ***D***. All experiments were performed in triplicate.

In order to fully characterize the two mutants, the knowledge of their colocalization frequency and T_{on}s as well as their basal phosphorylation is imperative. The SPT and WB data presented in ***Figure 7-2*** have been acquired by Dr L.C. Zanetti-Domingues (Zanetti-Domingues *et al.*, manuscript in preparation). The first two lines in plots ***Figure 7-2*** (**A**) and (**B**) show the results from basal and EGF-treated wt EGFR expressing cells (obtained from ***Figure 6-10***) for comparison.

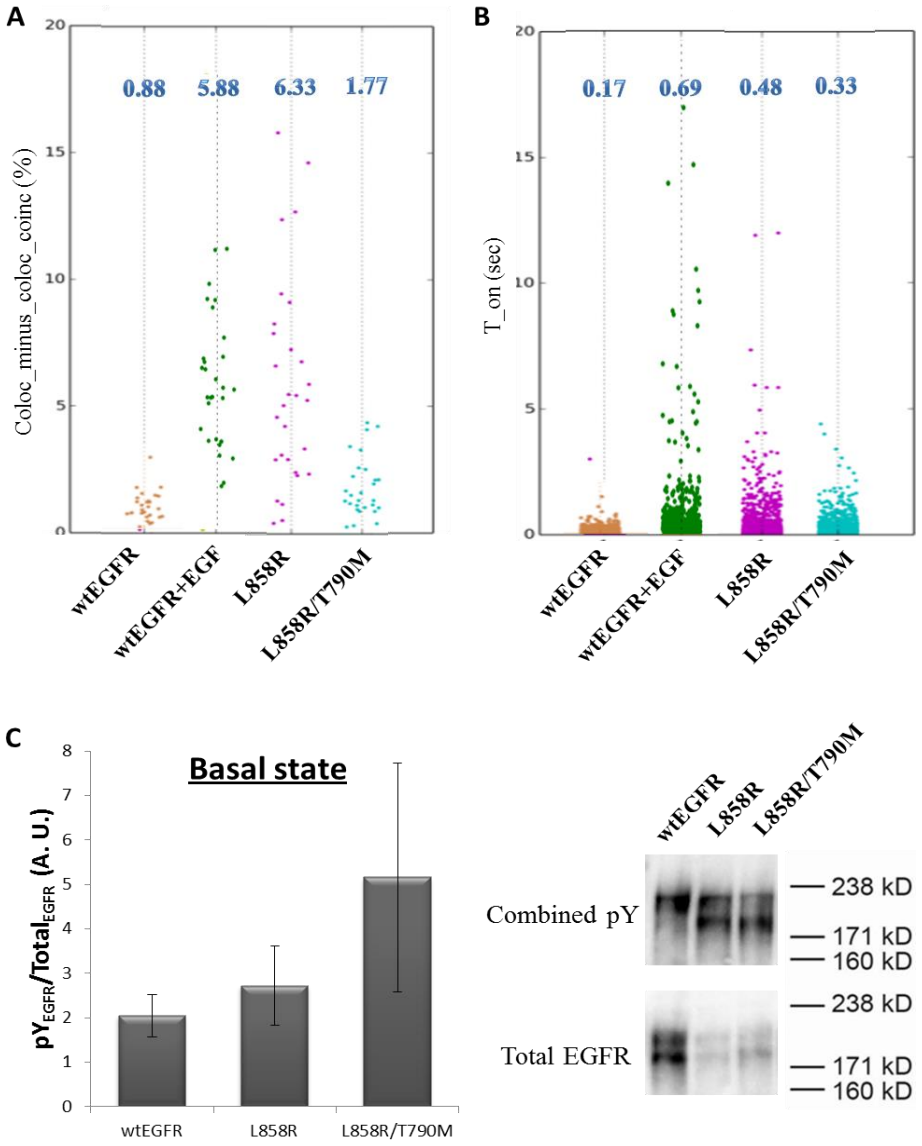


Figure 7-2: SPT analysis and receptor phosphorylation of the basal L858R and L858R/T790M EGFR expressing cells (adapted from Zanetti-Domingues *et al.*, manuscript in preparation). Cells expressing either L858R or L858R/T790M were prepared for both SPT and WB according to the general protocol described herein (***Sections 6.2.1*** and ***5.3.1*** respectively). Analysis of the SPT data was performed as in ***Section 6.2.1***. 1st and 2nd dataset in figures (**A**) and (**B**) show the results from experiments in CHO cells expressing wtEGFR with or without EGF obtained from ***Figure 6-10***.

L858R mutant not only exhibited a significant increase in the colocalization fraction compared to basal wt EGFR (1st line), consistent with the results from Valley *et al.*[103], but its colocalized fraction was comparable with the EGF-stimulated wt receptor (***Figure 7-2A***). The double mutant also showed a 2-fold increase in the formation of non-monomeric complexes (similarly to 9G8-treated wt receptors in ***Figure 6-17***) however its mean value was 3.5-fold lower compared to the single mutant. This last observation was unexpected as the additional mutation (T790M) on a L858R receptor has been suggested to enhance stability of the active state[281], a finding which is also in agreement with biochemical studies showing that the double mutant has higher basal phosphorylation than the single mutant[282]. In order to ensure that both mutants expressed in wt CHO cells show the same trend in phosphorylation as previously reported, receptor phosphorylation was also determined (also performed by Dr L.C. Zanetti-Domingues). WB analysis confirmed that the double mutant is hyper-phosphorylated in the absence of ligand (***Figure 7-2C***). Collectively these results appear to be consistent with what was suggested by the experiments in ***Chapter 6***: basal phosphorylation of the unliganded receptor is not correlated with a high degree of dimerization/oligomerization as opposed to what happens with the liganded receptor when treated with EGF concentrations close to physiological levels[234,251,283]. Despite the significant difference in the colocalization frequency between the two mutants, T_{on} of L858R was only 1.5-fold higher than that of L858R/T790M, a finding which may provide an explanation for the increased phosphorylation.

In order to better understand the effect of these mutations on the overall receptor geometry and ECD conformation, I performed FLImP and DOCA experiments. FLImP analysis demonstrated that L858R is highly oligomeric whereas the double mutant is mostly dimeric (***Figure 7-3 A*** and ***B*** respectively). These results are also consistent with the SPT results (***Figure 7-2***). Attempts to acquire good quality FLIM images were unsuccessful due to the low expression of the NSCLC mutants in CHO cells and hence low number of photons per pixel. However, Valley *et al.*[103] recently showed that L858R mutation favours the extended ECD conformation which is consistent, based on the model proposed herein, with its high oligomeric fingerprint (***Figure 7-3A***). Unfortunately there are no published data regarding the ECD conformation of the L858R/T790M mutant. With regards to the tethered-extended dimer ratio which was

shown in ***Chapter 6*** to be correlated with the phosphorylation of the unliganded receptor, L858R data (***Figure 7-3C***) suggest that Affibody separations consistent with the tethered dimer conformation are slightly favoured. This is in agreement with its basal phosphorylation status, slightly higher than basal wt EGFR, and the model in ***Figure 6-22***. On the other hand, the tethered-extended dimer ratio of L858R/T790M was unexpected. Even though, based on the results in ***Section 6.4*** (promotion of the tethered conformation in the basal receptor is accompanied by an increase in phosphorylation), its relatively low non-monomeric fraction, dimeric geometry and high basal phosphorylation led me to the assumption that the double mutant would preferably form tethered dimers, FLImP analysis revealed that there is not any such preference (***Figure 7-3C***).

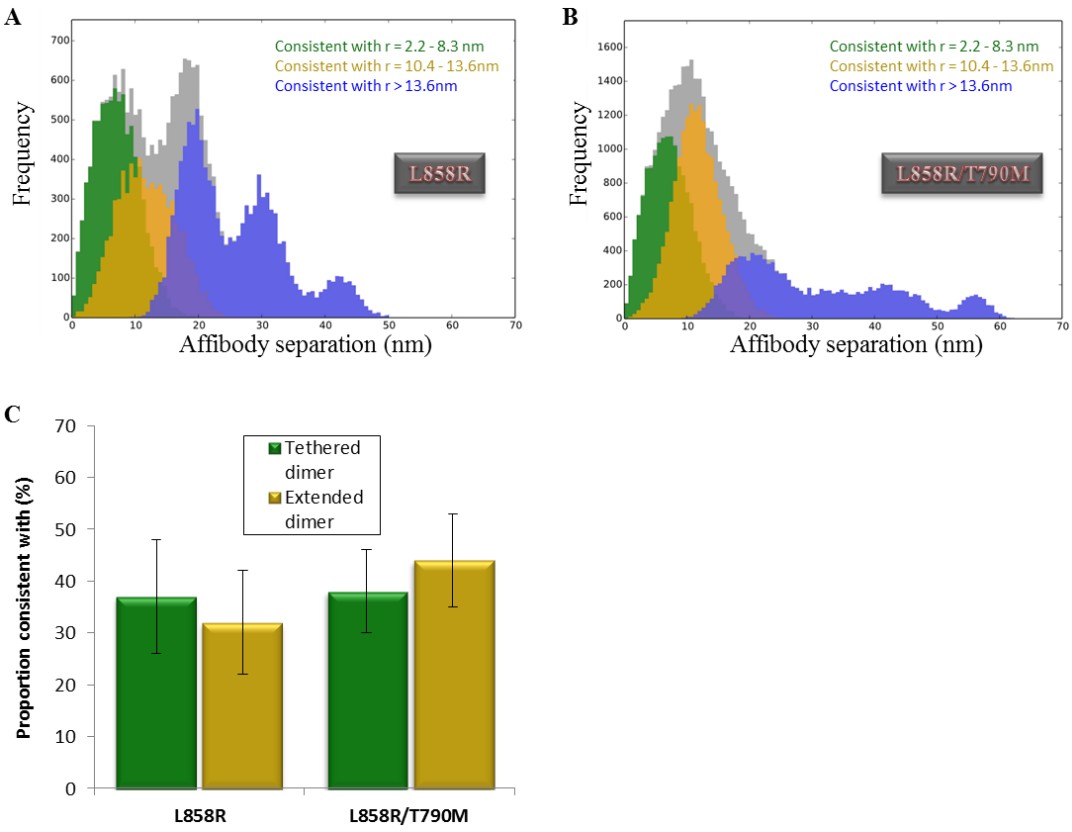


Figure 7-3: Determination of the geometry of NSCLC mutants in cells. FLImP results (**A** & **B**) of wt CHO cells expressing L858R or L858R/T790M EGFR mutants. All experiments were performed in triplicate. The relative populations of tethered and extended dimers of both mutants are shown in (**C**).

The reason why the proposed model presented herein failed to explain L858R/T790M hyperactivation may lie on a distinct characteristic of this mutant. Two independent studies have shown that a receptor harbouring both point mutations is still

phosphorylated even after disruption of the asymmetric dimer interface (either N- or C-lobe mutation)[282,284]. On the other hand, receptors with either the L858R or T790M are entirely dimerization-dependent. Furthermore, Jia *et al.*[285] recently showed that cells transfected with the activator-impaired version of the L858R/T790M were still able to proliferate. Therefore, the double mutation not only stabilizes the active conformation but most importantly, as the present data suggest, by-passes the activation mechanism followed by both the wt receptor and the constitutively active L858R mutant.

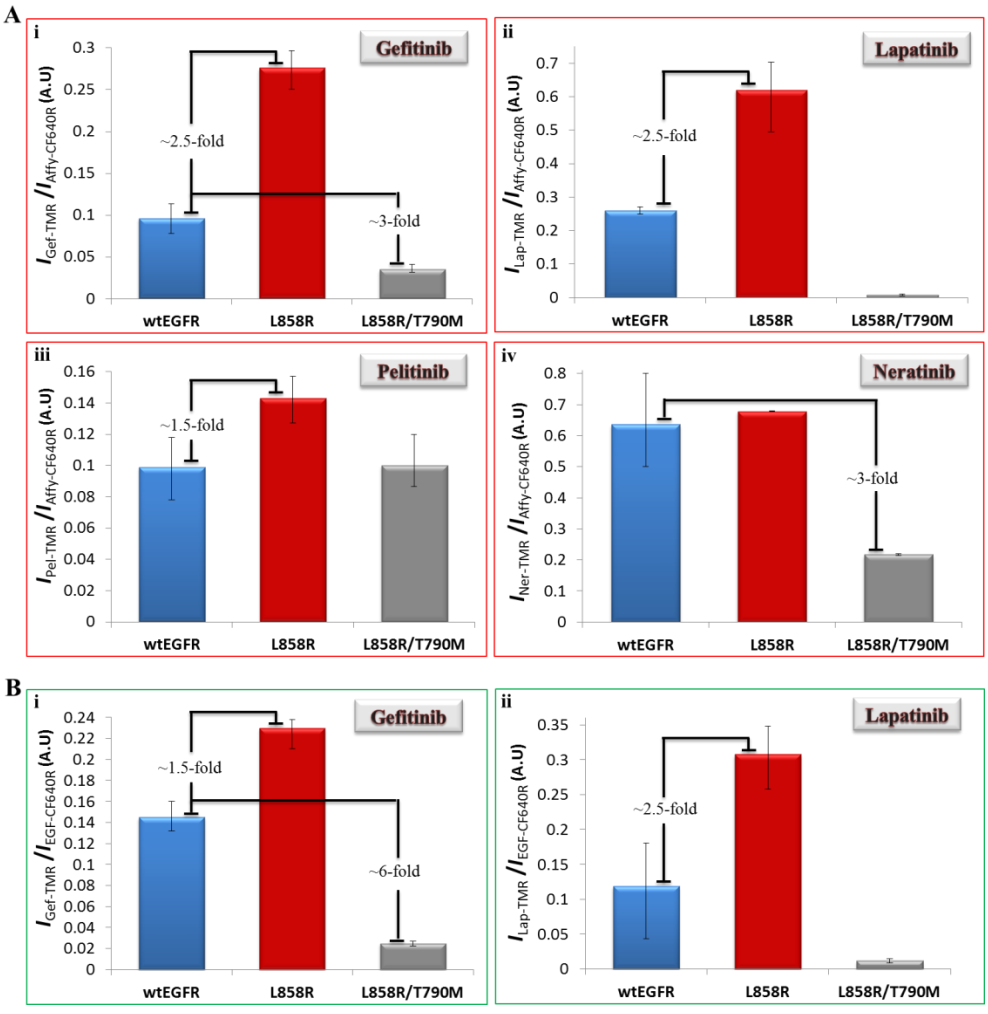
7.3 Drug Sensitivity/Resistance Mechanism of NSCLC EGFR Mutants

The T790M mutation was first identified as a secondary mutation, occurring in association with L858R in patients who had been independently treated with two reversible type I TKIs (Gefitinib[286] and Erlotinib[287]). While L858R conferred sensitivity to all type I TKIs[288], the so-called "gatekeeper" mutation, which occurs in many kinases[211], conferred resistant to both type I and II reversible TKIs[213]. It was initially though that the resistance mechanism of the T790M relies on the exclusion of the TKIs from the ATP site due to steric hindrance from the bulky side chain of methionine[212,286,287]. However, Yun *et al.*[213] showed that the resistance of this double mutant against competitive (reversible) TKIs (which retained low-nanomolar affinity) is due to increased affinity to ATP. On the other hand, irreversible TKIs effectively inhibit phosphorylation of the double mutant[212]. With the fTKIs in hand I sought to further explore the mechanism by which this mutant confers resistance to TKIs.

All the experiments and analysis were performed as described earlier (***Section 5.3.1***). The results from experiments in basal and EGF-treated cells expressing either L858R or L858R/T790M are shown in ***Figure 7-4***. For L858R expressing cells (red columns), I found that the membrane signal of both reversible and irreversible type I fTKIs was higher compared to wtEGFR in both basal and EGF-treated states (***Figure 7-4A & B, i*** and ***iii***). As discussed earlier (***Section 5.3.1***), any change in f-Pel (irreversible inhibitor) binding can only be attributed to alterations in its affinity, therefore the observed ~1.5-fold increase in Pelitinib binding (in both states) must be due to increased affinity as a result of the TKD point mutation. The fact that the structure and binding mode of Pelitinib are very similar to Gefitinib (***Figure 5-15***) suggests that the increased f-Gef

membrane signal is mainly due to increased affinity as a result of the L858R mutation (part of the increase in the basal state could be correlated to the slightly elevated basal phosphorylation). Collectively, the data indicate that the reason behind the increased affinity of type I TKIs against L858R EGFR mutant is the altered K_{on}.

While the higher affinity of type I TKIs found herein is consistent with previous reports[176,182,216], the increase in f-Lap binding (***Figure 7-4A & B, ii***) was unexpected. The fact that f-Ner binding (***Figure 7-4A & B, iv***) remained unaffected (which indicates that the overall number of inactive kinases has not changed compared to wt EGFR expressing cells) suggests that this increase can only be explained by increased Lapatinib affinity due to its slower dissociation rate (K_{off}) from the pocket. Furthermore, the type II fTKI data (and in agreement with the WB analysis in ***Figure 7-2C***) indicate that the L858R mutation does not itself promote the active conformation by disrupting the auto-inhibited interaction of the inactive state[49,278].



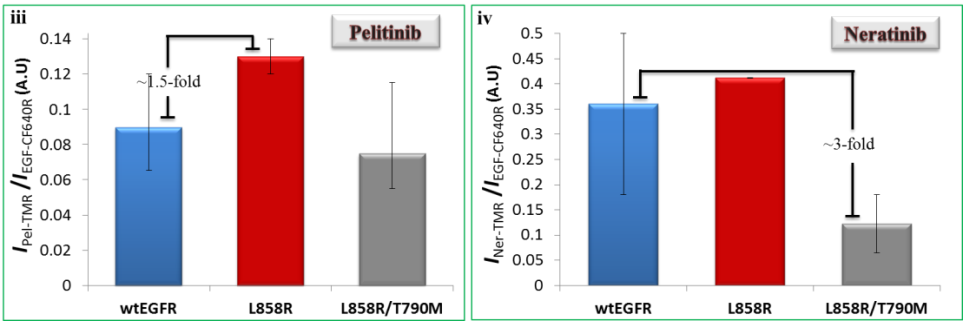


Figure 7-4: Determination of the resistance mechanism of NSCLC EGFR mutants. wt CHO cells expressing either L858R or L858R/T790M were prepared and labelled as described earlier for both the basal (A) and EGF-stimulated (B) receptors. All experiments were performed in triplicate.

This surprising result is in contrast with previous binding studies (where either purified EGFR TKD[176,182] or purified near full-length receptor[216] were used) which showed that Lapatinib binding affinity was decreased as a result of the L858R mutation. A L858R crystal structure resolved by Yoshikawa *et al.*[278] revealed a novel conformation where the ATP cleft is wider and more accessible, which is related to K_{on} of TKIs, than in the wt receptor. Even though this observation could explain the increased affinity of type I fTKIs against the L858R mutant, as the authors also suggested[278], the present type II fTKI data (their K_{on} is not affected) suggest that the increased Lapatinib affinity is independent of the reported altered feature of the ATP cleft, unless the observed widening also affected the dissociation rate (K_{off}) of type II fTKIs. Even though the exact mechanism behind the increased Lapatinib affinity is not understood and requires further studies, I speculate that it might be related either to the increased number of non-monomer complexes (Arkhipov *et al.* postulated that the ATP cleft of monomeric receptor is occluded[63]) found by SPT and FLImP (***Figure 7-2*** and ***Figure 7-3*** respectively) or to an alternative intermediate conformation (that is almost absent in wt receptor) which was shown by Sutto *et al.* to be the most stable in the L858R mutant (proposed by massive molecular dynamics simulation)[281].

Taken together the fTKI data presented herein (***Figure 7-4***) and the western blot analysis obtained from Dr L.C. Zanetti-Domingues (***Figure 7-2C***) for L858R suggest that the sensitization of the L858R EGFR to type I TKIs[25,276] is not exclusively due to constitutive activation, and consequently to increased affinity of type I TKIs, of the receptor by the single-point mutation, as previously thought[49]. First of all, the present study demonstrated that the L858R mutation in cells is not itself capable of hyper-

activating the receptor as Zhang *et al.* proposed[49]. My observation is also in agreement with a model recently suggested by Red Brewer *et al.*[282] (superacceptor hypothesis) who found that lung cancer mutants are more potent when they are co-expressed with the wt receptor and forced to adopt the acceptor position, a hypothesis which was also confirmed by Ruan *et al.* for the R776H EGFR mutant[289]. Secondly, the type II reversible TKI Lapatinib was also found to possess increased affinity against L858R and not reduced. Thus, the mechanism by which cancer cell lines harbouring L858R mutation confers sensitivity to type I TKIs may lie in the hyper-activation of the mutant (and the reduction of the number of inactive kinases) by another EGFR allele (superacceptor model). A similar sensitivity to type I inhibitors has also been shown to occur in the EGFR-overexpressing cancer cell line A431 in the basal state despite the fact that both type I and II TKIs have similar potencies for the wt receptor (which is the only EGFR allele expressed on A431 cells)[290]. By analogy to L858R EGFR expressing cancer cell lines, this sensitivity to type I TKIs is likely due to constitutive basal phosphorylation of the receptor by an autocrine loop in which cells self-provide the receptors with endogenous TGF-α (EGFR ligand)[291,292].

Next, the drug resistance mechanism of the L858R/T790M receptor was examined. Both type I and II reversible fTKIs (f-Gef and f-Lap) exhibited a decrease in their binding (***Figure 7-4A & B, i*** and ***ii***). In particular f-Lap binding was completely abolished as expected for a constitutive active receptor and also consistent with previous binding studies[176,182]. On the other hand, from the two irreversible fTKIs (f-Pel and f-Ner) only the type II f-Ner displayed a decrease in its binding, whereas f-Pel binding remained constant (***Figure 7-4A & B, iii*** and ***iv***). The fact that f-Ner still binds the receptors suggests that the observed reduction of the type II fTKIs is due to both the high number of active kinases as determined by WB (***Figure 7-2C***) and their previously reported reduced binding affinity[176,182]. On the contrary, the reduced f-Gef binding can only be associated to a reduction of its binding affinity, since f-Pel membrane signal is unaffected, and more specifically due to increased dissociation rate of the inhibitor from the pocket.

Yun *et al.*[213] demonstrated that the acquired drug resistance of the double mutant to reversible inhibitors[293,294] is due to the higher affinity to ATP and not due to steric hindrance from to the bulky methionine residue (which would affect substrate binding affinity). Even though in the context of the fTKI experiments the role of the competition

of the reversible TKIs with ATP cannot be examined, the present data show for the first time, and in contrast to Yun *et al.*[213], that the "gatekeeper" mutation also facilitates the release of the reversible inhibitors from the pocket. Therefore the role of the secondary T790M mutation in the drug resistance mechanism of the receptor is dual; assisting the faster release of TKIs from the pocket and also increasing the affinity for ATP. Furthermore the data suggest that irreversible type I TKIs would more effectively inhibit the double mutant than type II inhibitors.

8 General Conclusions and Future Perspective

This thesis describes the step-by-step development of novel chemical tools for the specific labelling of the TKD (intracellular) of ErbB family members and their utilization in fluorescent microscopy which aims to investigate, in combination with existing methods[70,71,136,137,251], the drugs mechanism of action and the basal architecture of EGFR on cells.

In order to specifically label the TKD of ErbB family members, EGFR TKIs, which bind to the ATP site of the receptors[22], were chosen to be turned into fluorescent imaging probes. The design of the fTKIs was based on the general structure of activity-based probes[122]. The synthesis of fTKIs (***Chapter 4***) was carried out using approved drugs or TKIs in clinical trials with good selectivity profiles[182,295] in order to minimize the off-target binding in the cellular environment. The final list of selected compounds for development of their fluorescent versions includes four TKIs: Gefitinib (type I reversible), Pelitinib (type I irreversible), Lapatinib (type II reversible) and Neratinib (type II irreversible). The diversity of the different classes of these TKIs allowed me to get an insight into their mechanism of action as well as to be able to use them, when combined, as reporters of EGFR activity in cells (***Chapters 5-7***). The synthesis of every compound reported herein (including the dyes) was based on previously reported synthesis. Even though some of the literature procedures did not work as expected, alternative reactions led to the desirable final compounds. Overall 15 final fluorescent EGFR inhibitors were synthesised herein.

Characterization of the affinity and selectivity of the synthesized fTKIs in a biochemical assay (KinomeScan®) demonstrated that the initial design and syntheses were successful as all compounds retained low nanomolar affinity and selectivity (***Table 5-1***). Furthermore, properties such as permeability and *in vitro* binding and specificity were assessed in cells (***Section 5.2***) and revealed:

i. the lipophilicity (ClogP) and the net charge of the dyes, as well as the size of the pharmacophore, affect the permeability of small-molecule conjugates

ii. different dyes can have different effects on the intracellular distribution and/or the non-specific binding of a conjugate

iii. reversible fTKIs cannot be used for live cell imaging

iv. both reversible and irreversible fTKIs can be used for labelling the receptors in pre-fixed cells

The observations yielded from my characterization could further provide an additional guidance for low-MW weight probe development in the future, especially for use in fluorescence imaging.

A key feature of the post-fixation fTKI labelling method is that it could allow one to directly monitor the binding affinity of different inhibitors to the receptor without perturbing its conformation (which can be induced by TKIs binding[65,217,219]). By exploiting this cell-based assay and combining it with quantitative imaging analysis (***Section 5.3.1***), I sought to get an insight into the conformationally selectivity of TKIs for the different EGFR TKD conformations. Even though the knowledge of such a preference could help the clinical need for better response of targeted therapeutics, the question regarding the selectivity of EGFR TKIs remained unanswered[218]. My study revealed that TKIs are indeed conformationally selective but also gave me an insight into the mechanism for this selectivity (binding kinetics). These findings highlight that the clinical success of a drug does not only depend on its binding affinity, determined by biochemical assays, but also on the type of the inhibitor (I or II) and the state of the TKD of the receptor. Furthermore, the fact that fTKIs do have preference for the different TKD conformations means they can be used as reporters of the receptor conformation.

Applications of fTKIs in different fluorescence microscopy techniques were also explored (***Sections 5.3.2*** and ***5.3.3***). These applications include intra-dimer FRET measurements, using FLIM-FRET, for assessing receptor dimerization and SPT using TIRF microscopy for studying receptor perturbation at single molecule level. Even though the preliminary results from both methods shown in ***Sections 5.3.2*** and ***5.3.3*** are promising, further optimization and studies are required in order to assess their validity. The information which could be potentially extracted from these methods could allow: 1) a better understanding into the mechanism of activation of EGFR-driven cancers without EGFR mutations as these cell types are likely to depend on homo- and hetero-dimerization[230] and 2) insight into dynamics of receptor inhibition and drug resistance. With regards to FLIM-FRET experiments, studies in stimulated cells where, presumably, the dimerization is higher is imperative for validating the method. SPT experiments require improvement of the SNR, by reducing the non-specific binding of

the fTKIs on the glass. In order to overcome this issue, alternative methods for loading the cells could be attempted[247–249]. Alternatively, different glass surface coating methods, such as polyelectrolyte multilayer coating, could reduce the non-specific binding on the glass[296].

In ***Chapters 6***, I employed a combinatorial approach in order to investigate the architecture of the basal full-length EGFR in its native membrane environment. While the ECD and TKD conformations of both the monomeric (inactive) and the EGF-bound (activated) dimeric receptor have been more extensively studied, the conformation of the pre-formed dimers is unknown. Determining their structure is of importance as it would allow scientists to build a model for the inactive-active transition which in turn would lead to the design of novel therapeutics agents.

The combinatorial approach included previously developed methods for probing the EGFR ECD structure[70,71,136,137] and combined with the pre-fixed fTKI labelling method reported herein for probing the TKD structure. The data from the study presented herein appears to be consistent with the model of the basal receptor illustrated in ***Figure 6-22***. According to that model (which does not include higher order oligomers), basal EGFR on the cell surface can adopt two different dimeric populations (apart from the inactive momoner[36]): inactive dimers where the ECDs adopt the extended conformation and the TKDs are in a symmetric configuration and active dimers with tethered ECDs coupled to asymmetric TKD dimers. Furthermore, the data indicated that oligomerization of the receptor in the basal state is not important for basal phosphorylation.

These findings further led me to investigate whether the mechanism of activation of two constitutively active NSCLC EGFR mutants, L858R and L858R/T790M, are consistent with the proposed model as well as the mechanism of drug resistance of these mutants (***Chapter 7***). In terms of the first objective, the data indicated that L858R EGFR mechanism of activation appears to be consistent with the model of the wt receptor (ligand-dependent activation) when it is not co-expressed with wt EGFR[282] (or perhaps other ErbB members). On the other hand, the double mutant (L858R/T790M) is able to "by-pass" the requirement for ligand binding as its kinase domain preferentially adopts the active conformation and, as recent studies indicated[282,284,285], do not require dimerization. With regards to the drug resistance mechanism, the data from the present study revealed:

- the sensitivity of L858R EGFR to type I TKIs is not only due to conformational change by the TKD mutation but also depends on the ability of the mutant to be hyper-phosphorylated by the wt receptor (super-acceptor hypothesis[282])
- the resistance of L858R/T790M to reversible TKIs is a result of both the higher affinity of the mutant for ATP[213] and the higher dissociation rate that reversible TKIs exhibit due to the presence of the bulky methionine.

In addition to this, fTKI data suggested that type I irreversible TKIs could more be more effective against such EGFR mutants as they bind both the inactive and the active TKD conformations with the same affinity.

The use of small-molecule fluorescence probes in fluorescence microscopy has proven a very challenging task, especially at single molecule level, due to the presence of the dyes. The properties of such low MW conjugates greatly depend on the net charge, lipophilicity and size of the dye. While neutral dyes are essential for the permeability of small molecule probes, they introduce artefacts at single molecule level due to their non-specific binding on the glass[161]. While labelling of pre-fixed cells can be performed at RT with a short incubation time (3-5 min) at low nM concentrations, live cell labelling (as in SPT) requires 2 hours incubation at 37^{o}C and a relatively high concentration. Even though different methods for reducing the non-specific binding were attempted (glass cleaning and photobleaching strategy) the SNR was not at the desired levels. Optimization of the glass coating method or use of alternative methods for loading cells with the probes could allow the use of fTKIs in techniques such as SPT and FLImP *etc.* in order to fully characterize the ICD of EGFR.

EGFR regulation has been also shown to be associated not only with the other ErbB family members[1,297] but also with other proteins such as c-MET[298] and G protein-coupled receptors[299]. Even though fTKIs were successfully used for studying the mechanism of action of EGFR drugs in CHO cells, expanding this method in other cell lines (which either express wt or other cancer EGFR mutants) would allow painting a complete picture of the TKIs mechanism. The method could be further exploited to investigate the binding of ErbB2 TKIs in cells. Both Lapatinib and Neratinib are known to inhibit ErbB2[300]. Therefore, expanding the method to ErbB2 would allow us get a better understanding of the ErbB2 mechanism of resistance either by mutation[301] or

other defence mechanism[302,303] as well as to study the effect of different dimerization partners in TKIs binding.

In addition to this, another future plan is the investigation of the effect of antibody treatment on the binding of different TKI types. Combinatorial treatment of both ECD and TKD inhibitors has been suggested as an alternative approach in clinics[18–21]. While EGFR inhibition by TKIs occurs by blocking ATP binding, the inhibition mechanism of antibodies varies[17,24,267]. Most importantly, it is not known whether antibody binding induces conformational changes of the TKD (as shown in ***Section 6.4***, 9G8 appears to promote asymmetric dimerization) since methods for probing the TKD conformation are not available. If such changes take place and since TKIs affinity depends on the TKD conformation (***Section 5.3.1***), the efficiency of antibody-TKI synergy could be affected by antibody binding. Therefore, exploring the effect of antibodies to the conformation of the TKD could potential help us to pre-determine which antibody-TKI combinations will exhibit the best synergistic effect.

Following validation of the intradimer FRET method and optimization of the single molecule methods (and in particular FLImP), the plans for the long term include the expansion of these methods to the other three members of the ErbB family. It has been suggested that ErbB family members not only interact with each other when activated[1,297] but also that they heretodimerize in the basal state[66–68,225]. The synthesised fTKIs to that can bind all four members in cells with low nanomolar affinity (***Table 5-1*** and ***Figure 5-10***). Therefore, fTKIs could in principle enable us to study the extent of dimerization when cells express all four members, the heterodimerization preference and to characterize the geometry and stoichiometry of ErbB family members in cells.

9 References

1. Lemmon, M. A. & Schlessinger, J. Cell signaling by receptor tyrosine kinases. *Cell* **141,** 1117–1134 (2010).

2. Ward, C. W., Lawrence, M. C., Streltsov, V. A., Adams, T. E. & McKern, N. M. The insulin and EGF receptor structures: new insights into ligand-induced receptor activation. *Trends Biochem. Sci.* **32,** 129–137 (2007).

3. Clayton, A. H. A. *et al.* Ligand-induced dimer-tetramer transition during the activation of the cell surface epidermal growth factor receptor-A multidimensional microscopy analysis. *J. Biol. Chem.* **280,** 30392–30399 (2005).

4. Clayton, A. H. A., Orchard, S. G., Nice, E. C., Posner, R. G. & Burgess, A. W. Predominance of activated EGFR higher-order oligomers on the cell surface. *Growth Factors* **26,** 316–324 (2008).

5. Cohen, S. Isolation of a mouse submaxillary gland protein accelerating incisor eruption and eyelid opening in the new-born animal. *J. Biol. Chem.* **237,** 1555–1562 (1962).

6. Cohen, S. & Levi-Montalcini, R. Purification and properties of a nerve growth-promoting factor isolated from mouse sarcoma 180. *Cancer Res.* **17,** 15–20 (1957).

7. Cohen, S. The stimulation of epidermal proliferation by a specific protein (EGF). *Dev. Biol.* **12,** 394–407 (1965).

8. Cohen, S., Carpenter, G. & King, L. J. Epidermal growth factor-receptor-protein kinase interactions. *J. Biol. Chem.* **225,** 4834–4842 (1980).

9. Yarden, Y. & Pines, G. The ERBB network: at last, cancer therapy meets systems biology. *Nat. Rev. Cancer* **12,** 553–563 (2012).

10. Yarden, Y. & Sliwkowski, M. X. Untangling the ErbB signalling network. *Nat. Rev. Mol. Cell Biol.* **2,** 127–137 (2001).

11. Wilson, K. J., Gilmore, J. L., Foley, J., Lemmon, M. A. & Riese, D. J. Functional Selectivity of EGF Family Peptide Growth Factors: Implications for Cancer. *Pharmacol. Ther.* **122,** 1–8 (2009).

12. Shi, F., Telesco, S. E., Liu, Y., Radhakrishnan, R. & Lemmon, M. a.

ErbB3/HER3 intracellular domain is competent to bind ATP and catalyze autophosphorylation. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 7692–7697 (2010).

13. Wagner, M. J., Stacey, M. M., Liu, B. A. & Pawson, T. Molecular mechanisms of SH2- and PTB-Domain-containing proteins in receptor tyrosine kinase signaling. *Cold Spring Harb. Perspect. Biol.* **5,** (2013).

14. Thompson, D. M. & Gill, G. N. The EGF receptor: structure, regulation and potential role in malignancy. *Cancer Surv* **4,** 767–788 (1985).

15. Roskoski, R. The ErbB/HER family of protein-tyrosine kinases and cancer. *Pharmacol. Res.* **79,** 34–74 (2014).

16. Ranson, M. Epidermal growth factor receptor tyrosine kinase inhibitors. *Br J Cancer* **90,** 2250–2255 (2004).

17. Martinelli, E., De Palma, R., Orditura, M., De Vita, F. & Ciardiello, F. Anti-epidermal growth factor receptor monoclonal antibodies in cancer therapy. *Clin. Exp. Immunol.* **158,** 1–9 (2009).

18. Noto, A. *et al.* Combination therapy with anti-ErbB3 monoclonal antibodies and EGFR TKIs potently inhibits non-small cell lung cancer. *Oncotarget* **4,** 1253–65 (2013).

19. Leung, W. *et al.* Combining lapatinib and pertuzumab to overcome lapatinib resistance due to NRG1-mediated signalling in HER2-amplified breast cancer. *Oncotarget* **6,** 5678–94 (2015).

20. Huang, S., Armstrong, E. a, Benavente, S., Chinnaiyan, P. & Harari, P. M. Dual-Agent Molecular Targeting of the Epidermal Growth Factor Receptor (EGFR): Combining Anti-EGFR Antibody with Tyrosine Kinase Inhibitor Dual-Agent Molecular Targeting of the Epidermal Growth Factor Receptor (EGFR): Combining Anti-EGFR Antibody with. *Cancer Res.* **64,** 5355–5362 (2004).

21. Weickhardt, a. J. *et al.* Dual Targeting of the Epidermal Growth Factor Receptor Using the Combination of Cetuximab and Erlotinib: Preclinical Evaluation and Results of the Phase II DUX Study in Chemotherapy-Refractory, Advanced Colorectal Cancer. *J. Clin. Oncol.* **30,** 1505–1512 (2012).

22. Zhang, J., Yang, P. L. & Gray, N. S. Targeting cancer with small molecule kinase inhibitors. *Nat Rev Cancer* **9,** 28–39 (2009).

23. Toporkiewicz, M., Meissner, J., Matusewicz, L., Czogalla, A. & Sikorski, A. F. Toward a magic or imaginary bullet? Ligands for drug targeting to cancer cells: principles, hopes, and challenges. *Int J Nanomedice* **10,** 1399–1414 (2015).

24. Schmiedel, J., Blaukat, A., Li, S., Knoechel, T. & Ferguson, K. M. Matuzumab binding to EGFR prevents the conformational rearrangement required for dimerization. *Cancer Cell* **13,** 365–373 (2009).

25. Yun, C. H. *et al.* Structures of Lung Cancer-Derived EGFR Mutants and Inhibitor Complexes: Mechanism of Activation and Insights into Differential Inhibitor Sensitivity. *Cancer Cell* **11,** 217–227 (2007).

26. Greulich, H. *et al.* Functional analysis of receptor tyrosine kinase mutations in lung cancer identifies oncogenic extracellular domain mutations of ERBB2. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 14476–81 (2012).

27. Umelo, I. *et al.* Identification of a novel HER3 activating mutation homologous to EGFR-L858R in lung cancer. *Oncotarget* **7,** (2015).

28. Soung, Y. H. *et al.* Somatic mutations of the ERBB4 kinase domain in human cancers. *Int. J. Cancer* **118,** 1426–1429 (2006).

29. Sergina, N. V *et al.* Escape from HER-family tyrosine kinase inhibitor therapy by the kinase-inactive HER3. *Nature* **445,** 437–441 (2007).

30. Abd El-Rehim, D. M. *et al.* Expression and co-expression of the members of the epidermal growth factor receptor (EGFR) family in invasive breast carcinoma. *Br. J. Cancer* **91,** 1532–42 (2004).

31. Wong, A. J. *et al.* Structural alterations of the epidermal growth factor receptor gene in human gliomas. *Proc. Natl. Acad. Sci. U. S. A.* **89,** 2965–9 (1992).

32. Zhan, L., Xiang, B. & Muthuswamy, S. K. Controlled activation of ErbB1/ErbB2 heterodimers promote invasion of three-dimensional organized epithelia in an ErbB1-dependent manner: Implications for progression of ErbB2-overexpressing tumors. *Cancer Res.* **66,** 5201–5208 (2006).

33. Lemmon, M. A. *et al.* Two EGF molecules contribute additively to stabilization of the EGFR dimer. *EMBO J.* **16,** 281–294 (1997).

34. Garrett, T. P. J. *et al.* Crystal structure of a truncated epidermal growth factor

receptor extracellular domain bound to transforming growth factor alpha. *Cell* **110,** 763–73 (2002).

35. Ogiso, H. Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. *Cell* **110,** 775–787 (2002).

36. Ferguson, K. M. *et al.* EGF Activates Its Receptor by Removing Interactions that Autoinhibit Ectodomain Dimerization. *Mol. Cell* **11,** 507–517 (2003).

37. Cho, H.-S. & Leahy, D. J. Structure of the extracellular region of HER3 reveals an interdomain tether. *Science* **297,** 1330–1333 (2002).

38. Bouyain, S., Longo, P. A., Li, S., Ferguson, K. M. & Leahy, D. J. The extracellular region of ErbB4 adopts a tethered conformation in the absence of ligand. *Proc. Natl. Acad. Sci. U.S.A.* **102,** 15024–15029 (2005).

39. Alvarado, D., Klein, D. E. & Lemmon, M. a. ErbB2 resembles an autoinhibited invertebrate epidermal growth factor receptor. *Nature* **461,** 287–291 (2009).

40. Garrett, T. P. J. *et al.* The crystal structure of a truncated ErbB2 ectodomain reveals an active conformation, poised to interact with other ErbB receptors. *Mol. Cell* **11,** 495–505 (2003).

41. Cho, H.-S. *et al.* Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. *Nature* **421,** 756–760 (2003).

42. Liu, P. *et al.* A single ligand is sufficient to activate EGFR dimers. *Proc. Natl. Acad. Sci.* **109,** 10861–10866 (2012).

43. Gotoh, N., Tojo, A., Hino, M., Yazaki, Y. & Shibuya, M. A highly conserved tyrosine residue at codon 845 within the kinase domain is not required for the transforming activity of human epidermal growth factor receptor. *Biochem. Biophys. Res. Commun.* **186,** 768–774 (1992).

44. Mueller, K. L., Powell, K., Madden, J. M., Eblen, S. T. & Boerner, J. L. EGFR Tyrosine 845 Phosphorylation-Dependent Proliferation and Transformation of Breast Cancer Cells Require Activation of p38 MAPK. *Transl. Oncol.* **5,** 327–34 (2012).

45. Boerner, J. L., Demory, M. L., Silva, C. & Parsons, S. J. Phosphorylation of Y845 on the Epidermal Growth Factor Receptor Mediates Binding to the

Mitochondrial Protein Cytochrome c Oxidase Subunit II Phosphorylation of Y845 on the Epidermal Growth Factor Receptor Mediates Binding to the Mitochondrial Protein Cyt. *Mol Cell Biol* **24,** 7059–7071 (2004).

46. Stamos, J., Sliwkowski, M. X. & Eigenbrot, C. Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. *J. Biol. Chem.* **277,** 46265–46272 (2002).

47. Wood, E. R. A Unique Structure for Epidermal Growth Factor Receptor Bound to GW572016 (Lapatinib): Relationships among Protein Conformation, Inhibitor Off-Rate, and Receptor Activity in Tumor Cells. *Cancer Res.* **64,** 6652–6659 (2004).

48. Ferguson, K. M. A structure-based view of Epidermal Growth Factor Receptor regulation. *Annu. Rev. Biophys.* **37,** 353–373 (2008).

49. Zhang, X., Gureasko, J., Shen, K., Cole, P. A. & Kuriyan, J. An Allosteric Mechanism for Activation of the Kinase Domain of Epidermal Growth Factor Receptor. *Cell* **125,** 1137–1149 (2006).

50. Zhang, X. *et al.* Inhibition of the EGF receptor by binding of MIG6 to an activating kinase domain interface. *Nature* **450,** 741–744 (2007).

51. Qiu, C. *et al.* Mechanism of Activation and Inhibition of the HER4/ErbB4 Kinase. *Structure* **16,** 460–467 (2008).

52. Monsey, J., Shen, W., Schlesinger, P. & Bose, R. Her4 and Her2/neu tyrosine kinase domains dimerize and activate in a reconstituted in vitro system. *J. Biol. Chem.* **285,** 7035–7044 (2010).

53. Brewer, M. R. *et al.* The Juxtamembrane Region of the EGFR functions as an activation domain. *Mol. Cell* **34,** 641–651 (2009).

54. Jura, N. *et al.* Mechanism for Activation of the EGF Receptor Catalytic Domain by the Juxtamembrane Segment. *Cell* **137,** 1293–1307 (2009).

55. Aertgeerts, K. *et al.* Structural analysis of the mechanism of inhibition and allosteric activation of the kinase domain of HER2 protein. *J. Biol. Chem.* **286,** 18756–18765 (2011).

56. Jura, N., Shan, Y., Cao, X., Shaw, D. E. & Kuriyan, J. Structural analysis of the

catalytically inactive kinase domain of the human EGF receptor 3. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 21608–21613 (2009).

57. Ishikawa, T. *et al.* Design and synthesis of novel human epidermal growth factor receptor 2 (HER2)/epidermal growth factor receptor (EGFR) dual inhibitors bearing a pyrrolo[3,2-d]pyrimidine scaffold. *J. Med. Chem.* **54,** 8030–8050 (2011).

58. Thiel, K. W. & Carpenter, G. Epidermal growth factor receptor juxtamembrane region regulates allosteric tyrosine kinase activation. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 19238–19243 (2007).

59. Sakisaka, T., Itoh, T., Miura, K. & Takenawa, T. Phosphatidylinositol 4,5-bisphosphate phosphatase regulates the rearrangement of actin filaments. *Mol Cell Biol* **17,** 3841–3849 (1997).

60. Michailidis, I. E. *et al.* Phosphatidylinositol-4,5-bisphosphate regulates epidermal growth factor receptor activation. *Pflugers Arch. Eur. J. Physiol.* **461,** 387–397 (2011).

61. Halim, K. B. A., Kolds??, H. & Sansom, M. S. P. Interactions of the EGFR juxtamembrane domain with PIP2-containing lipid bilayers: Insights from multiscale molecular dynamics simulations. *Biochim. Biophys. Acta - Gen. Subj.* **1850,** 1017–1025 (2015).

62. Endres, N. F. *et al.* Conformational coupling across the plasma membrane in activation of the EGF receptor. *Cell* **152,** 543–556 (2013).

63. Arkhipov, A. *et al.* Architecture and membrane interactions of the EGF receptor. *Cell* **152,** 557–569 (2013).

64. Doerner, A., Scheck, R. & Schepartz, A. Growth Factor Identity Is Encoded by Discrete Coiled-Coil Rotamers in the EGFR Juxtamembrane Region. *Chem. Biol.* **22,** 776–784 (2015).

65. Mi, L.-Z. *et al.* Simultaneous visualization of the extracellular and cytoplasmic domains of the epidermal growth factor receptor. *Nat. Struct. Mol. Biol.* **18,** 984–989 (2011).

66. Moriki, T., Maruyama, H. & Maruyama, I. N. Activation of preformed EGF receptor dimers by ligand-induced rotation of the transmembrane domain. *J. Mol.*

Biol. **311,** 1011–1026 (2001).

67. Landau, M., Fleishman, S. J. & Ben-Tal, N. A putative mechanism for downregulation of the catalytic activity of the EGF receptor via direct contact between its kinase and C-terminal domains. *Structure* **12,** 2265–2276 (2004).

68. Tao, R.-H. & Maruyama, I. N. All EGF(ErbB) receptors have preformed homo- and heterodimeric structures in living cells. *J. Cell Sci.* **121,** 3207–3217 (2008).

69. Martin-Fernandez, M., Clarke, D. T., Tobin, M. J., Jones, S. V & Jones, G. R. Preformed oligomeric epidermal growth factor receptors undergo an ectodomain structure change during signaling. *Biophys. J.* **82,** 2415–27 (2002).

70. Needham, S. R. *et al.* Measuring EGFR Separations on Cells with ~10 nm Resolution via Fluorophore Localization Imaging with Photobleaching. *PLoS One* **8,** 1–13 (2013).

71. Zanetti-Domingues, L. C. *et al.* Determining the geometry of oligomers of the human epidermal growth factor family on cells with 7 nm resolution. *Prog. Biophys. Mol. Biol.* **118,** 139–152 (2015).

72. Nevoltris, D. & Chames, P. Tethered Epidermal Growth Factor Receptor Involved in EGFR / ErbB2. 1388–1399 (2015).

73. Webb, S. E. D. *et al.* Single-molecule imaging and fluorescence lifetime imaging microscopy show different structures for high- and low-affinity epidermal growth factor receptors in A431 cells. *Biophys. J.* **94,** 803–819 (2008).

74. Zhang, Q., Park, E., Kani, K. & Landgraf, R. Functional isolation of activated and unilaterally phosphorylated heterodimers of ERBB2 and ERBB3 as scaffolds in ligand-dependent signaling. *Proc. Natl. Acad. Sci.* **109,** 13237–13242 (2012).

75. Bragin, P. E. *et al.* HER2 Transmembrane Domain Dimerization Coupled with Self-Association of Membrane-Embedded Cytoplasmic Juxtamembrane Regions. *J. Mol. Biol.* **428,** 52–61 (2016).

76. Bocharov, E. V. *et al.* Spatial structure of the dimeric transmembrane domain of the growth factor receptor ErbB2 presumably corresponding to the receptor active state. *J. Biol. Chem.* **283,** 6950–6956 (2008).

77. Mineev, K. S. *et al.* Spatial Structure of the Transmembrane Domain

Heterodimer of ErbB1 and ErbB2 Receptor Tyrosine Kinases. *J. Mol. Biol.* **400,** 231–243 (2010).

78. Mineev, K. S. *et al.* Spatial structure and dimer-monomer equilibrium of the ErbB3 transmembrane domain in DPC micelles. *Biochim. Biophys. Acta - Biomembr.* **1808,** 2081–2088 (2011).

79. Bocharov, E. V., Mineev, K. S., Goncharuk, M. V. & Arseniev, A. S. Structural and thermodynamic insight into the process of 'weak' dimerization of the ErbB4 transmembrane domain by solution NMR. *Biochim. Biophys. Acta - Biomembr.* **1818,** 2158–2170 (2012).

80. Matsushita, C. *et al.* Transmembrane helix orientation influences membrane binding of the intracellular juxtamembrane domain in Neu receptor peptides. *Proc. …* **110,** 1646–51 (2013).

81. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy Principles of Fluorescence Spectroscopy*. *Principles of fluorescence spectroscopy, Springer, New York, USA, 3rd edn, 2006.* (2006). doi:10.1007/978-0-387-46312-4

82. Stender AS, Marchuk K, Liu C, Sander S, Meyer MW, Smith EA, Neupane B, Wang G, Li J, Cheng JX, Huang B, Fang N. Single cell optical imaging and spectroscopy. *Chem. Rev.* **113,** 2469–527 (2013).

83. Ishikawa-Ankerhold, H. C., Ankerhold, R. & Drummen, G. P. C. Advanced fluorescence microscopy techniques-FRAP, FLIP, FLAP, FRET and FLIM. *Molecules* **17,** 4047–4132 (2012).

84. Truong, K. & Mitsuhiko, I. The use of FRET imaging microscopy to detect protein – protein interactions and protein conformational changes in vivo. *Curr. Opin. Struct. Biol.* **11,** 573–578 (2001).

85. Swift, S. R. & Trinkle-Mulcahy, L. Basic principles of FRAP , FLIM and FRET. *Proc. R. Microsc. Soc.* **39,** 3–10 (2004).

86. Becker, W. Fluorescence lifetime imaging - techniques and applications. *J. Microsc.* **247,** 119–136 (2012).

87. Wallrabe, H. & Periasamy, A. Imaging protein molecules using FRET and FLIM microscopy. *Curr. Opin. Biotechnol.* **16,** 19–27 (2005).

88. Martin-Fernandez, M. L. Human epidermal growth factor receptor (HER1) aligned on the plasma membrane adopts key features of Drosophila EGFR asymmetry. *Biochem. Soc. Trans.* **40,** 184–8 (2012).

89. Sauer, M., Hofkens, J. & Enderlein, J. Basic Principles of Fluorescence Spectroscopy. *Handb. Fluoresc. Spectrosc. Imaging From Single Mol. to Ensembles* 1–30 (2011). doi:10.1002/9783527633500.ch1

90. Joo, C., Balci, H., Ishitsuka, Y., Buranachai, C. & Ha, T. Advances in single-molecule fluorescence methods for molecular biology. *Annu. Rev. Biochem.* **77,** 51–76 (2008).

91. Funatsu, T., Harada, Y., Tokunaga, M., Saito, K. & Yanagida, T. Imaging of single fluorescent molecules and individual ATP turnovers by single myosin molecules in aqueous solution. *Nature* **374,** 555–559 (1995).

92. Petty, H. R. Fluorescence Microscopy: Established and Emerging Methods, Experimental Strategies, and Applications in Immunology. *Microsc. Res. Tech.* **70,** 687–709 (2007).

93. Sako, Y. & Yanagida, T. Single-molecule visualization in cell biology. *Nat. Rev. Mol. cell Biol.* **Suppl,** SS1–5 (2003).

94. Zlatanova, J. & van Holde, K. Single-Molecule Biology: What Is It and How Does It Work? *Mol. Cell* **24,** 317–329 (2006).

95. Kapanidis, A. N. & Strick, T. Biology, one molecule at a time. *Trends Biochem. Sci.* **34,** 234–243 (2009).

96. Rolfe, D. J. *et al.* Automated multidimensional single molecule fluorescence microscopy feature detection and tracking. *Eur. Biophys. J.* **40,** 1167–1186 (2011).

97. Colomb, W. & Sarkar, S. K. Extracting physics of life at the molecular level: A review of single-molecule data analyses. *Phys. Life Rev.* **13,** 107–137 (2015).

98. Alcor, D., Gouzer, G. & Triller, A. Single-particle tracking methods for the study of membrane receptors dynamics. *Eur. J. Neurosci.* **30,** 987–997 (2009).

99. D. Axelrod, D.E. Koppel, J. Schelessinger, E. Elson, W. W. W. Mobility Measurement by Analysis of Fluorescence Photobeaching recovery Kinetics.

Biophys. J. **16,** 1055–1069 (1976).

100. Magde, D., Elson, E. & Webb, W. Thermodynamic Fluctuations in a Reacting System— Measurement by Fluorescence Correlation Spectroscopy. *Phys. Rev. Lett.* **29,** 705–708 (1972).

101. Sako, Y., Minoghchi, S. & Yanagida, T. Single-molecule imaging of EGFR signalling on the surface of living cells. *Nat. Cell Biol.* **2,** 168–172 (2000).

102. Chung, I. *et al.* Spatial control of EGF receptor activation by reversible dimerization on living cells. *Nature* **464,** 783–787 (2010).

103. Valley, C. C. *et al.* Enhanced dimerization drives ligand-independent activity of mutant EGFR in lung cancer. *Mol. Biol. Cell* **13,** 854 – 865 (2015).

104. Zanetti-Domingues, L. C., Martin-Fernandez, M. L., Needham, S. R., Rolfe, D. J. & Clarke, D. T. A Systematic Investigation of Differential Effects of Cell Culture Substrates on the Extent of Artifacts in Single-Molecule Tracking. *PLoS One* **7,** (2012).

105. Betzig, E. *et al.* Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313,** 1642–1645 (2006).

106. Rust, M. J., Bates, M. & Zhuang, X. W. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat Methods* **3,** 793–795 (2006).

107. HELL, S. W. & WICHMANN, J. Breaking the Diffraction Resolution Limit By Stimulated-Emission - Stimulated-Emission-Depletion Fluorescence Microscopy. *Opt. Lett.* **19,** 780–782 (1994).

108. Tetin, S. Y. & Stroupe, S. D. Antibodies in diagnostic applications. *Curr. Pharm. Biotechnol.* **5,** 9–16 (2004).

109. Binz, H. K., Amstutz, P. & Pluckthun, A. Engineering novel binding proteins from nonimmunoglobulin domains. *Nat. Biotechnol.* **23,** 1257–1268 (2005).

110. Paduch, M. *et al.* Generating conformation-specific synthetic antibodies to trap proteins in selected functional states. *Methods* **60,** 3–14 (2013).

111. Rothbauer, U. *et al.* Targeting and tracing antigens in live cells with fluorescent nanobodies. *Nat. Methods* **3,** 887–9 (2006).

112. Maier, J., Traenkle, B. & Rothbauer, U. Real-time analysis of epithelial-mesenchymal transition using fluorescent single-domain antibodies. *Sci. Rep.* **5,** 13402 (2015).

113. Wojcik, J. *et al.* Allosteric Inhibition of Bcr-Abl Kinase by High-Affinity Monobody Inhibitors Directed to the SH2-Kinase Interface. *J. Biol. Chem.* **291,** 8836–8847 (2016).

114. Best, M. D. Click chemistry and bioorthogonal reactions: Unprecedented selectivity in the labeling of biological molecules. *Biochemistry* **48,** 6571–6584 (2009).

115. De Graaf, A. J., Kooijman, M., Hennink, W. E. & Mastrobattista, E. Nonnatural amino acids for site-specific protein conjugation. *Bioconjug. Chem.* **20,** 1281–1295 (2009).

116. Sunbul, M. & Yin, J. Site specific protein labeling by enzymatic posttranslational modification. *Org. Biomol. Chem.* **7,** 3361–3371 (2009).

117. Young, T. S. & Schultz, P. G. Beyond the canonical 20 amino acids: Expanding the genetic lexicon. *J. Biol. Chem.* **285,** 11039–11044 (2010).

118. Chen, Y. X., Triola, G. & Waldmann, H. Bioorthogonal chemistry for site-specific labeling and surface immobilization of proteins. *Acc. Chem. Res.* **44,** 762–773 (2011).

119. Schmied, W. H., Els??sser, S. J., Uttamapinant, C. & Chin, J. W. Efficient multisite unnatural amino acid incorporation in mammalian cells via optimized pyrrolysyl tRNA synthetase/tRNA expression and engineered eRF1. *J. Am. Chem. Soc.* **136,** 15577–15583 (2014).

120. Yano, Y. & Matsuzaki, K. Tag-probe labeling methods for live-cell imaging of membrane proteins. - Supplementary. *Biochim. Biophys. Acta* **1788,** 2124–2131 (2009).

121. Dillingham, M. S. & Wallace, M. I. Protein modification for single molecule fluorescence microscopy. *Org. Biomol. Chem.* **6,** 3031–3037 (2008).

122. Cravatt, B. F., Wright, A. T. & Kozarich, J. W. Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu. Rev. Biochem.* **77,** 383–414 (2008).

123. Liu, Y., Patricelli, M. P. & Cravatt, B. F. Activity-based protein profiling: the serine hydrolases. *Proc. Natl. Acad. Sci. U. S. A.* **96,** 14694–9 (1999).

124. Blair, J. A. *et al.* Structure-guided development of affinity probes for tyrosine kinases using chemical genetics. *Nat. Chem. Biol.* **3,** 229–38 (2007).

125. Zhang, Z. *et al.* Leveraging kinase inhibitors to develop small molecule tools for imaging kinases by fluorescence microscopy. *Mol. Biosyst.* **8,** 2523–2526 (2012).

126. Thurber, G. M. *et al.* Single-cell and subcellular pharmacokinetic imaging allows insight into drug action in vivo. *Nat. Commun.* **4,** 1504 (2013).

127. Lanning, B. R. *et al.* A road map to evaluate the proteome-wide selectivity of covalent kinase inhibitors. *Nat. Chem. Biol.* **10,** 1–10 (2014).

128. Devaraj, N. K., Hilderbrand, S., Upadhyay, R., Mazitschek, R. & Weissleder, R. Bioorthogonal turn-on probes for imaging small molecules inside living cells. *Angew. Chemie - Int. Ed.* **49,** 2869–2872 (2010).

129. Cohen, M. S., Hadjivassiliou, H. & Taunton, J. A clickable inhibitor reveals context-dependent autoactivation of p90 RSK. *Nat. Chem. Biol.* **3,** 156–160 (2007).

130. Bunnage, M. E., Gilbert, A. M., Jones, L. H. & Hett, E. C. Know your target, know your molecule. *Nat. Chem. Biol.* **11,** 368–372 (2015).

131. Niphakis, M. J. & Cravatt, B. F. Enzyme inhibitor discovery by activity-based protein profiling. *Annu. Rev. Biochem.* **83,** 341–77 (2014).

132. Barkovich, K. J. *et al.* Kinetics of inhibitor cycling underlie therapeutic disparities between EGFR-driven lung and brain cancers. *Cancer Discov.* **2,** 450–457 (2012).

133. Edgington, L. E. *et al.* Functional imaging of legumain in cancer using a new quenched activity-based probe. *J. Am. Chem. Soc.* **135,** 174–182 (2013).

134. Uddin, M. J. *et al.* Selective visualization of cyclooxygenase-2 in inflammation and cancer by targeted fluorescent imaging agents. *Cancer Res.* **70,** 3618–3627 (2010).

135. Uddin, M. J., Crews, B. C., Ghebreselasie, K. & Marnett, L. J. Design, synthesis, and structure-activity relationship studies of fluorescent inhibitors of

cycloxygenase-2 as targeted optical imaging agents. *Bioconjugate Chem.* **24,** 712–23 (2013).

136. Tynan, C. J. *et al.* Human Epidermal Growth Factor Receptor (EGFR) Aligned on the Plasma Membrane Adopts Key Features of Drosophila EGFR Asymmetry. *Mol. Cell. Biol.* **31,** 2241–2252 (2011).

137. Coban, O. *et al.* Effect of phosphorylation on EGFR dimer stability probed by single-molecule dynamics and FRET/FLIM. *Biophys. J.* **108,** 1013–1026 (2015).

138. Wissner, A. *et al.* Synthesis and structure-activity relationships of 6,7-disubstituted 4-anilinoquinoline-3-carbonitriles. The design of an orally active, irreversible inhibitor of the tyrosine kinase activity of the epidermal growth factor receptor (EGFR) and the human epi. *J Med Chem* **46,** 49–63 (2003).

139. Seimbille, Y., Phelps, M. E., Czernin, J. & Silverman, D. H. S. Fluorine-18 labeling of 6,7-disubstituted anilinoquinazoline derivatives for positron emission tomography (PET) imaging of tyrosine kinase receptors: Synthesis of18F-Iressa and related molecular probes. *J. Label. Compd. Radiopharm.* **48,** 829–843 (2005).

140. Wang, J.-Q., Gao, M., Miller, K. D., Sledge, G. W. & Zheng, Q.-H. Synthesis of [11C]Iressa as a new potential PET cancer imaging agent for epidermal growth factor receptor tyrosine kinase. *Bioorg. Med. Chem. Lett.* **16,** 4102–4106 (2006).

141. Fang, Z. *et al.* Lithium chloride-catalyzed selective demethylation of aryl methyl ethers under microwave irradiation. *J. Mol. Catal. A Chem.* **274,** 16–23 (2007).

142. Harris, C. S., Kettle, J. G. & Williams, E. J. Facile synthesis of 7-amino anilinoquinazolines via direct amination of the quinazoline core. *Tetrahedron Lett.* **46,** 7381–7384 (2005).

143. Shreder, K. R., Wong, M. S., Nomanbhoy, T., Leventhal, P. S. & Fuller, S. R. Synthesis of AX7593, a quinazoline-derived photoaffinity probe for EGFR. *Org. Lett.* **6,** 3715–3718 (2004).

144. Lee, Y. B. & Ahn, C. H. Quinazoline derivatives and therapeutic us thereof. *US Pat., 20050187231 A1* (2005).

145. Mahboobi, S. *et al.* Novel chimeric histone deacetylase inhibitors: A series of lapatinib hybrides as potent inhibitors of epidermal growth factor receptor

(EGFR), human epidermal growth factor receptor 2 (HER2), and histone deacetylase activity. *J. Med. Chem.* **53,** 8546–8555 (2010).

146. Abraham, S. *et al.* Raf kinase modulator compounds and methods of use thereof. *US Pat., 20110118245 A1* (2011).

147. Tung, R. 4-aminoquinazoline derivatives and methods of use thereof. *US Pat., 20110097320 A1* (2011). doi:US 20100322867A1

148. Patel, G. *et al.* Kinase scaffold repurposing for neglected disease drug discovery: discovery of an efficacious, lapatinib-derived lead compound for trypanosomiasis. *J Med Chem* **56,** 3820–3832 (2013).

149. Roschangar, F., Brown, J. C., Cooley, B. E., Sharp, M. J. & Matsuoka, R. T. Use of lithium N,O-dimethylhydroxylamide as an efficient in situ protecting agent for aromatic aldehydes. *Tetrahedron* **58,** 1657–1666 (2002).

150. Seto, M. & Ohashi, T. Fused nitrogen-comprising heterocycle compound. *US Pat., 201002345351* (2010).

151. Kirincich, S. J. *et al.* Benzhydrylquinazolinediones: Novel cytosolic phospholipase A2α inhibitors with improved physicochemical properties. *Bioorganic Med. Chem.* **17,** 4383–4405 (2009).

152. Strohmeier, G. A., Fabian, W. M. F. & Uray, G. A Combined Experimental and Theoretical Approach toward the Development of Optimized Luminescent Carbostyrils. *Helv.Chim.Acta* **87,** 215–226 (2004).

153. Tsou, H. R. ., Overbeek-Klumpers, E. & Wissner, A. Tricyclic protein kinase inhibitors. *US Pat., 20030065180 A1* (2003). doi:US 20100322867A1

154. Carmi, C. *et al.* Irreversible Inhibition of Epidermal Growth Factor Receptor Activity by 3-Aminopropanamides. *J Med Chem* **55,** 2251–2264 (2012).

155. Spicka, K. J. Design and synthesis of fluorescent dyes for use in proteomic research. (Montana State University).

156. Gießler, K., Griesser, H., Göhringer, D., Sabirov, T. & Richert, C. Synthesis of 3′-BODIPY-labeled active esters of nucleotides and a chemical primer extension assay on beads. *European J. Org. Chem.* 3611–3620 (2010). doi:10.1002/ejoc.201000210

157. Dohi, T., Morimoto, K., Maruyama, A. & Kita, Y. Direct synthesis of bipyrroles using phenyliodine bis(trifluoroacetate) with bromotrimethylsilane. *Org. Lett.* **8,** 2007–2010 (2006).

158. Lee, J. S. *et al.* Synthesis of a BODIPY library and its application to the development of live cell glucagon imaging probe. *J. Am. Chem. Soc.* **131,** 10077–10082 (2009).

159. Uddin, M. J. & Marnett, L. J. Synthesis of 5- and 6-Carboxy-X-rhodamines. *Org. Lett.* **10,** 4799–4801 (2008).

160. Yu, H., Xiao, Y. & Guo, H. From spirolactam mixtures to regioisomerically pure 5- and 6-rhodamines: A chemodosimeter-Inspired Strategy. *Org. Lett.* **14,** 2014–2017 (2012).

161. Zanetti-Domingues, L. C., Tynan, C. J., Rolfe, D. J., Clarke, D. T. & Martin-Fernandez, M. Hydrophobic Fluorescent Probes Introduce Artifacts into Single Molecule Tracking Experiments Due to Non-Specific Binding. *PLoS One* **8,** (2013).

162. Roovers, R. C. *et al.* A biparatopic anti-EGFR nanobody efficiently inhibits solid tumour growth. *Int. J. Cancer* **129,** 2013–2024 (2011).

163. Schneider, C. a, Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9,** 671–675 (2012).

164. Hall, M. *et al.* The WEKA data mining software. *ACM SIGKDD Explor. Newsl.* **11,** 10 (2009).

165. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9,** 676–682 (2012).

166. Webb, S. E. D., Needham, S. R., Roberts, S. K. & Martin-Fernandez, M. L. Multidimensional single-molecule imaging in live cells using total-internal-reflection fluorescence microscopy. *Opt. Lett.* **31,** 2157 (2006).

167. Chen, W. S. *et al.* Requirement for intrinsic protein tyrosine kinase in the immediate and late actions of the EGF receptor. *Nature* **328,** 820–823 (1987).

168. Honegger, A. M. *et al.* Point mutation at the ATP binding site of EGF receptor abolishes protein-tyrosine kinase activity and alters cellular routing. *Cell* **51,**

199–209 (1987).

169. Levitzki, A. & Mishani, E. Tyrphostins and other tyrosine kinase inhibitors. *Annu. Rev. Biochem.* **75,** 93–109 (2006).

170. Yaish, P., Gazit, a, Gilon, C. & Levitzki, a. Blocking of EGF-dependent cell proliferation by EGF receptor kinase inhibitors. *Science* **242,** 933–935 (1988).

171. Zhou, W. *et al.* Novel mutant-selective EGFR kinase inhibitors against EGFR T790M. *Nature* **462,** 1070–1074 (2009).

172. Baselga, J. Targeting tyrosine kinases in cancer: the second wave. *Science* **312,** 1175–1178 (2006).

173. Tan, C.-S., Cho, B.-C. & Soo, R. A. Next-generation epidermal growth factor receptor tyrosine kinase inhibitors in epidermal growth factor receptor -mutant non-small cell lung cancer. *Lung Cancer* **93,** 59–68 (2016).

174. Roskoski, R. Classification of small molecule protein kinase inhibitors based upon the structures of their drug-enzyme complexes. *Pharmacol. Res.* **103,** 26–48 (2016).

175. Traxler, P. *et al.* AEE788: a dual family epidermal growth factor receptor/ErbB2 and vascular endothelial growth factor receptor tyrosine kinase inhibitor with antitumor and antiangiogenic activity. *Cancer Res.* **64,** 4931–4941 (2004).

176. Davis, M. I. *et al.* Comprehensive analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **29,** 1046–1051 (2011).

177. Moulder, S. L. *et al.* Abstract A143: ARRY-380, a selective HER2 inhibitor: From drug design to clinical evaluation. *Mol. Cancer Ther.* **10 ,** A143–A143 (2011).

178. Jani, J. P. *et al.* Discovery and pharmacologic characterization of CP-724,714, a selective ErbB2 tyrosine kinase inhibitor. *Cancer Res.* **67,** 9887–9893 (2007).

179. Li, D. *et al.* BIBW2992, an irreversible EGFR/HER2 inhibitor highly effective in preclinical lung cancer models. *Oncogene* **27,** 4702–4711 (2008).

180. Barlaam, B. *et al.* Discovery of AZD8931, an equipotent, reversible inhibitor of signaling by EGFR, HER2, and HER3 receptors. *ACS Med. Chem. Lett.* **4,** 742–746 (2013).

181. Tsou, H.-R. *et al.* Optimization of 6,7-disubstituted-4-(arylamino)quinoline-3-carbonitriles as orally active, irreversible inhibitors of human epidermal growth factor receptor-2 kinase activity. *J. Med. Chem.* **48,** 1107–1131 (2005).

182. Karaman, M. W. *et al.* A quantitative analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **26,** 127–32 (2008).

183. Fabian, M. A. *et al.* A small molecule-kinase interaction map for clinical kinase inhibitors. *Nat. Biotechnol.* **23,** 329–36 (2005).

184. Fry, D. W. Mechanism of action of erbB tyrosine kinase inhibitors. *Exp. Cell Res.* **284,** 131–139 (2003).

185. Wissner, A. & Mansour, T. S. The development of HKI-272 and related compounds for the treatment of cancer. *Arch. Pharm. Chem. Life Sci.* **341,** 465–477 (2008).

186. Peng, Y. H. *et al.* Protein kinase inhibitor design by targeting the Asp-Phe-Gly (DFG) motif: The role of the DFG motif in the design of epidermal growth factor receptor inhibitors. *J. Med. Chem.* **56,** 3889–3903 (2013).

187. Thurber, G. M., Reiner, T., Yang, K. S., Kohler, R. H. & Weissleder, R. Effect of Small-Molecule Modification on Single-Cell Pharmacokinetics of PARP Inhibitors. *Mol. Cancer Ther.* **13,** 986–995 (2014).

188. Kim, E. *et al.* Optimized Near-IR Fluorescent Agents for in Vivo Imaging of Btk Expression. *Bioconjug. Chem.* **26,** 1513–1518 (2015).

189. Murrey, H. E. *et al.* Systematic Evaluation of Bioorthogonal Reactions in Live Cells with Clickable HaloTag Ligands: Implications for Intracellular Imaging. *J. Am. Chem. Soc.* **137,** 11461–11475 (2015).

190. Dekker, F. J. *et al.* Small-molecule inhibition of APT1 affects Ras localization and signaling. *Nat. Chem. Biol.* **6,** 449–456 (2010).

191. Watanabe, S., Mizukami, S., Akimoto, Y., Hori, Y. & Kikuchi, K. Intracellular protein labeling with prodrug-like probes using a mutant ??-lactamase tag. *Chem. - A Eur. J.* **17,** 8342–8349 (2011).

192. Petrov, K. G. *et al.* Optimization and SAR for dual ErbB-1/ErbB-2 tyrosine kinase inhibition in the 6-furanylquinazoline series. *Bioorganic Med. Chem. Lett.*

16, 4686–4691 (2006).

193. Wissner, A. *et al.* 4-Anilino-6,7-Dialkoxyquinoline-3-Carbonitrile Inhibitors of Epidermal Growth Factor Receptor Kinase and Their Bioisosteric Relationship To the 4-Anilino-6,7-Dialkoxyquinazoline Inhibitors. *J. Med. Chem.* **43,** 3244–3256 (2000).

194. Rezende, L. C. D. D. L., Emery, S. & Emery, F. A review of the synthetic strategies for the development of BODIPY dyes for conjugation with proteins. *Orbital Elec. J. Chem* **5,** 62–83 (2013).

195. Meltola, N. J., Wahlroos, R. & Soini, A. E. Hydrophilic labeling reagents of dipyrrylmethene-BF2 dyes for two-photon excited fluorometry: Syntheses and photophysical characterization. *J. Fluoresc.* **14,** 635–647 (2004).

196. Kang, H. C. . & Haugland, R. P. Long wavelength chemically reactive dipyrromethene boron difluoride dyes and conjugates. *US Pat., 5274113* (1993).

197. Zhang, M. *et al.* Synthesis of pyrrolyldipyrrinato BF2 complexes by oxidative nucleophilic substitution of boron dipyrromethene with pyrrole. *Org. Biomol. Chem.* **10,** 2139 (2012).

198. Dale, C. L., Hill, S. J. & Kellam, B. New potent, short-linker BODIPY-630/650^{TM} labelled fluorescent adenosine receptor agonists. *Medchemcomm* **3,** 333 (2012).

199. Beija, M., Afonso, C. a M. & Martinho, J. M. G. Synthesis and applications of Rhodamine derivatives as fluorescent probes. *Chem. Soc. Rev.* **38,** 2410–2433 (2009).

200. Meyer, R. & Sundmacher, W. Zur Kenntniss des m-Amidophenols. *Ber. Dtsch. Chem. Ges.* **32,** 2112–2124 (1899).

201. Boyarskiy, V. P. *et al.* Photostable, amino reactive and water-soluble fluorescent labels based on sulfonated rhodamine with a rigidized xanthene fragment. *Chem. - A Eur. J.* **14,** 1784–1792 (2008).

202. Mitronova, G. Y. *et al.* New fluorinated rhodamines for optical microscopy and nanoscopy. *Chem. - A Eur. J.* **16,** 4477–4488 (2010).

203. Montalbetti, C. A. G. N. & Falque, V. Amide bond formation and peptide

coupling. *Tetrahedron* **61,** 10827–10852 (2005).

204. Kowada, T., Maeda, H. & Kikuchi, K. BODIPY-based probes for the fluorescence imaging of biomolecules in living cells. *Chem. Soc. Rev.* **44,** 4953–72 (2015).

205. Olivier, J. H., Widmaier, J. & Ziessel, R. Near-infrared fluorescent nanoparticles formed by self-assembly of lipidic (bodipy) dyes. *Chem. - A Eur. J.* **17,** 11709–11714 (2011).

206. Macdonald-Obermann, J. L. & Pike, L. J. The intracellular juxtamembrane domain of the epidermal growth factor (EGF) receptor is responsible for the allosteric regulation of EGF binding. *J. Biol. Chem.* **284,** 13570–13576 (2009).

207. Kronqvist, N. *et al.* Combining phage and staphylococcal surface display for generation of ErbB3-specific Affibody molecules. *Protein Eng. Des. Sel.* **24,** 385–396 (2011).

208. Liu, Y. & Gray, N. S. Rational design of inhibitors that bind to inactive kinase conformations. *Nat. Chem. Biol.* **2,** 358–364 (2006).

209. Park, J. H. & Lemmon, M. A. Occupy EGFR. *Cancer Discov.* **2,** 398–400 (2012).

210. Ravichandran, S., Luke, B. T. & Collins, J. R. Can structural features of kinase receptors provide clues on selectivity and inhibition? A molecular modeling study. *J. Mol. Graph. Model.* **57,** 36–48 (2015).

211. Zuccotto, F., Ardini, E., Casale, E. & Angiolini, M. Through the 'gatekeeper door': Exploiting the active kinase conformation. *J. Med. Chem.* **53,** 2681–2694 (2010).

212. Kwak, E. L. *et al.* Irreversible inhibitors of the EGF receptor may circumvent acquired resistance to gefitinib. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 7665–70 (2005).

213. Yun, C.-H. *et al.* The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 2070–5 (2008).

214. Solca, F. *et al.* Target binding properties and cellular activity of afatinib (BIBW 2992), an irreversible ErbB family blocker. *J. Pharmacol. Exp. Ther.* **343,** 342–

50 (2012).

215. Qiu, C. *et al.* In Vitro Enzymatic Characterization of Near Full Length EGFR in Activated and Inhibited States. *Biochemistry* **48,** 6624–6632 (2009).

216. Wang, Z. *et al.* Mechanistic insights into the activation of oncogenic forms of EGF receptor. *Nat. Struct. Mol. Biol.* **18,** 1388–93 (2011).

217. Lu, C., Mi, L. Z., Schürpf, T., Walz, T. & Springer, T. A. Mechanisms for kinase-mediated dimerization of the epidermal growth factor receptor. *J. Biol. Chem.* **287,** 38244–38253 (2012).

218. Park, J., Liu, Y., Lemmon, M. & Radhakrishnan, R. Erlotinib binds both inactive and active conformations of the EGFR tyrosine kinase domain. *Biochem. J.* **448,** 417–423 (2012).

219. Anido, J. *et al.* ZD1839 , a Specific Epidermal Growth Factor Receptor (EGFR) Tyrosine Kinase Inhibitor , Induces the Formation of Inactive EGFR / HER2 and EGFR / HER3 Heterodimers and Prevents Heregulin Signaling in HER2-overexpressing Breast. *Clin. cancer Res.* **9,** 1274–1283 (2003).

220. Hofman, E. G. *et al.* EGF induces coalescence of different lipid rafts. *J. Cell Sci.* **121,** 2519–2528 (2008).

221. Zhang, X., Crespo, A. & Fernández, A. Turning promiscuous kinase inhibitors into safer drugs. *Trends Biotechnol.* **26,** 295–301 (2008).

222. Gajiwala, K. S. *et al.* Insights into the aberrant activity of mutant EGFR kinase domain and drug recognition. *Structure* **21,** 209–219 (2013).

223. Gazdar, A. F. Activating and resistance mutations of EGFR in non-small-cell lung cancer: role in clinical response to EGFR tyrosine kinase inhibitors. *Oncogene* **28,** S24–S31 (2009).

224. Murray, C. W. & Rees, D. C. The rise of fragment-based drug discovery. *Nat. Chem.* **1,** 187–192 (2009).

225. Yu, X., Sharma, K. D., Takahashi, T., Iwamoto, R. & Mekada, E. Ligand-independent Dimer Formation of Epidermal Growth Factor Receptor (EGFR) Is a Step Separable from Ligand-induced EGFR Signaling Xiaochun. *Mol. Biol. Cell* **13,** 2547–2557 (2002).

226. Gadella, T. W. J. & Jovin, M. Oligomerization of Epidermal Growth Factor Receptors on A431 Cells Studied by Time-resolved Fluorescence Imaging Microscopy . A Stereochemical Model for Tyrosine Kinase Receptor Activation. *J. Biol. Chem.* **129,** 1543–1558 (1995).

227. Gan, H. K. *et al.* The epidermal growth factor receptor (EGFR) tyrosine kinase inhibitor AG1478 increases the formation of inactive untethered EGFR dimers: Implications for combination therapy with monoclonal antibody 806. *J. Biol. Chem.* **282,** 2840–2850 (2007).

228. Cochet, C. *et al.* Demonstration of epidermal growth factor-induced receptor dimerization in living cells using a chemical covalent cross-linking agent. *J. Biol. Chem.* **263,** 3290–5 (1988).

229. Tanner, K. G. & Kyte, J. Dimerization of the Extracellular Domain of the Receptor for Epidermal Growth Factor Containing the Membrane-spanning Segment in Response to Treatment with Epidermal Growth Factor *. *J. Biol. Chem.* **274,** 35985–35990 (1999).

230. Waterhouse, B. R. *et al.* Assessment of EGFR/HER2 dimerization by FRET-FLIM utilizing Alexa-conjugated secondary antibodies in relation to targeted therapies in cancers. *Oncotarget* **2,** 728–36 (2011).

231. Maruyama, I. N. Mechanisms of activation of receptor tyrosine kinases: monomers or dimers. *Cells* **3,** 304–30 (2014).

232. Blakely, B. T. *et al.* Epidermal growth factor receptor dimerization monitored in live cells. *Nat Biotech* **18,** 218–222 (2000).

233. Liu, P. *et al.* Investigation of the dimerization of proteins from the epidermal growth factor receptor family by single wavelength fluorescence cross-correlation spectroscopy. *Biophys. J.* **93,** 684–98 (2007).

234. Yamashita, H., Yano, Y., Kawano, K. & Matsuzaki, K. Oligomerization-function relationship of EGFR on living cells detected by the coiled-coil labeling and FRET microscopy. *Biochim. Biophys. Acta - Biomembr.* **1848,** 1359–1366 (2015).

235. Bublil, E. M. *et al.* Kinase-mediated quasi-dimers of EGFR. *FASEB J.* **24,** 4744–4755 (2010).

236. Szabó, Á., Horváth, G., Szöllősi, J. & Nagy, P. Quantitative Characterization of the Large-Scale Association of ErbB1 and ErbB2 by Flow Cytometric Homo-FRET Measurements. *Biophys. J.* **95,** 2086–2096 (2008).

237. Bader, A. N., Hofman, E. G., Voortman, J., Van Bergen En Henegouwen, P. M. P. & Gerritsen, H. C. Homo-FRET imaging enables quantification of protein cluster sizes with subcellular resolution. *Biophys. J.* **97,** 2613–2622 (2009).

238. Hofman, E. G. *et al.* Ligand-induced EGF receptor oligomerization is kinase-dependent and enhances internalization. *J. Biol. Chem.* **285,** 39481–39489 (2010).

239. Lax, I. *et al.* Epidermal growth factor (EGF) induces oligomerization of soluble, extracellular, ligand-binding domain of EGF receptor: A low resolution projection structure of the ligand-binding domain. *J. Biol. Chem.* **266,** 13828–13833 (1991).

240. *Molecular Probes Handbook: A Guide to Fluorescent Probes and Labeling Technologies, 11th Edition*. (2011).

241. Sailer, B. L., Steinkamp, J. A. & Crissman, H. A. Flow cytometric fluorescence lifetime analysis of DNA-binding probes. *Eur J Histochem.* **42 (Spec N,** 19–27 (1998).

242. Cui, H. H., Valdez, J. G., Steinkamp, J. A. & Crissman, H. A. Fluorescence lifetime discrimination of cellular DNA and RNA using various intercalating dyes and flow cytometric analysis. Optical Diagnostics of Living Cells IV. *Proc SPIE* **4260,** 175–183 (2001).

243. Van Zandvoort, M. A. M. J. *et al.* Discrimination of DNA and RNA in cells by a vital fluorescent probe: Lifetime imaging of SYTO13 in healthy and apoptotic cells. *Cytometry* **47,** 226–235 (2002).

244. Ramachandran, R., Tweten, R. K. & Johnson, A. E. The domains of a cholesterol-dependent cytolysin undergo a major FRET-detected rearrangement during pore formation. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 7139–44 (2005).

245. Deka, C. *et al.* Analysis of fluorescence lifetime and quenching of FITC-conjugated antibodies on cells by phase-sensitive flow cytometry. *Cytometry* **25,** 271–279 (1996).

246. Hern, J. A. *et al.* Formation and dissociation of M1 muscarinic receptor dimers seen by total internal reflection fluorescence imaging of single molecules. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 2693–8 (2010).

247. Mobbs, P., Becker, D., Williamson, R., Bate, M. & Warner, A. in *Ogden D (ed) The Plymouth workshop handbook. The Company of Biologists Ltd. Cambridge* 361–387 (1994).

248. Hovis, K. R., Padmanabhan, K. & Urban, N. N. A Simple Method of In Vitro Electroporation Allows Visualization, Recording, and Calcium Imaging of Local Neuronal Circuitslic Access. *J Neurisci Methods* **191,** 1–10 (2010).

249. Sharei, A. *et al.* A vector-free microfluidic platform for intracellular delivery. *Proc Natl Acad Sci U S A* **110,** 2082–2087 (2013).

250. Sun, C., Cao, Z., Wu, M. & Lu, C. Intracellular Tracking of Single Native Molecules with Electroporation-delivered Quantum Dots. *Anal. Chem.* **86,** 11403–11409 (2014).

251. Needham, S. R. *et al.* EGFR oligomerization organizes kinase-active dimers into competent signalling platforms. *Nat. Commun.* **7,** 13307 (2016).

252. Bessman, N. J., Bagchi, A., Ferguson, K. M. & Lemmon, M. A. Complex Relationship between Ligand Binding and Dimerization in the Epidermal Growth Factor Receptor. *Cell Rep.* **9,** 1306–1317 (2014).

253. Kozer, N. *et al.* Evidence for extended YFP-EGFR dimers in the absence of ligand on the surface of living cells. *Phys. Biol.* **8,** 066002 (2011).

254. Ziomkiewicz, I. *et al.* Dynamic conformational transitions of the EGF receptor in living mammalian cells determined by FRET and fluorescence lifetime imaging microscopy. *Cytom. Part A* **83,** 794–805 (2013).

255. Kaszuba, K. *et al.* N-Glycosylation as determinant of epidermal growth factor receptor conformation in membranes. *Proc. Natl. Acad. Sci. U. S. A.* **112,** 4334–4339 (2015).

256. Nordberg, E. *et al.* Cellular studies of binding, internalization and retention of a radiolabeled EGFR-binding affibody molecule. *Nucl. Med. Biol.* **34,** 609–618 (2007).

257. Nordberg, E. *et al.* Effects of an EGFR-binding affibody molecule on intracellular signaling pathways. *Ann. Intern. Med.* **36,** 967–972 (2010).

258. Kovacs, E. *et al.* Analysis of the role of the C-terminal tail in the regulation of the epidermal growth factor receptor. *Mol. Cell. Biol.* **1162,** MCB.00248–15 (2015).

259. Macdonald-Obermann, J. L., Piwnica-Worms, D. & Pike, L. J. Mechanics of EGF receptor/ErbB2 kinase activation revealed by luciferase fragment complementation imaging. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 137–42 (2012).

260. Sogabe, S. *et al.* Structure-Based Approach for the Discovery of Pyrrolo[3,2-d]pyrimidine-Based EGFR T790M/L858R Mutant Inhibitors. *ACS Med. Chem. Lett.* **4,** 201–205 (2013).

261. McLaughlin, S., Smith, S. O., Hayman, M. J. & Murray, D. An electrostatic engine model for autoinhibition and activation of the epidermal growth factor receptor (EGFR/ErbB) family. *J. Gen. Physiol.* **126,** 41–53 (2005).

262. Stateva, S. R. *et al.* The activating role of phospho-(Tyr)-calmodulin on the epidermal growth factor receptor. *Biochem J* **472,** 195–204 (2015).

263. Walker, F. *et al.* CR1/CR2 interactions modulate the functions of the cell surface epidermal growth factor receptor. *J. Biol. Chem.* **279,** 22387–22398 (2004).

264. Needham, S. R. *et al.* Determining the geometry of oligomers of the human epidermal growth factor family on cells with <10nm resolution. *Biochem. Soc. Trans.* **43,** 309–14 (2015).

265. Huang, Y. *et al.* Molecular basis for multimerization in the activation of the Epidermal Growth Factor Receptor. *Elife* **5,** 14107 (2016).

266. Stirling Churchman, L., Flyvbjerg, H. & Spudich, J. A. A Non-Gaussian Distribution Quantifies Distances Measured with Fluorescence Localization Techniques. *Biophys. J.* **90,** 668–671 (2006).

267. Li, S. *et al.* Structural basis for inhibition of the epidermal growth factor receptor by cetuximab. *Cancer Cell* **7,** 301–311 (2005).

268. Schmitz, K. R., Bagchi, A., Roovers, R. C., Van Bergen En Henegouwen, P. M. P. & Ferguson, K. M. Structural evaluation of EGFR inhibition mechanisms for nanobodies/VHH domains. *Structure* **21,** 1214–1224 (2013).

269. Littlefield, P. *et al.* Structural analysis of the EGFR/HER3 heterodimer reveals the molecular basis for activating HER3 mutations. *Sci. Signal.* **7,** ra114 (2014).

270. Björkelund, H., Gedda, L., Barta, P., Malmqvist, M. & Andersson, K. Gefitinib induces epidermal growth factor receptor dimers which alters the interaction characteristics with 125I-EGF. *PLoS One* **6,** (2011).

271. Greenall, S. a, Donoghue, J. F., Gottardo, N. G., Johns, T. G. & Adams, T. E. Glioma-specific Domain IV EGFR cysteine mutations promote ligand-induced covalent receptor dimerization and display enhanced sensitivity to dacomitinib in vivo. *Oncogene* **34,** 1–9 (2015).

272. Li, L. *et al.* An EGFR wild type-EGFRvIII-HB-EGF feed-forward loop regulates the activation of EGFRvIII. *Oncogene* **33,** 4253–64 (2014).

273. Montagut, C. *et al.* Identification of a mutation in the extracellular domain of the Epidermal Growth Factor Receptor conferring cetuximab resistance in colorectal cancer. *Nat. Med.* **18,** 221–3 (2012).

274. Massarelli, E., Johnson, F. M., Erickson, H. S., Wistuba, I. I. & Papadimitrakopoulou, V. Uncommon Epidermal Growth Factor Receptor mutations in non-small cell lung cancer and their mechanisms of EGFR tyrosine kinase inhibitors sensitivity and resistance. *Lung Cancer* **80,** 235–241 (2013).

275. Park, A. K. J., Francis, J. M., Park, W.-Y., Park, J.-O. & Cho, J. Constitutive asymmetric dimerization drives oncogenic activation of epidermal growth factor receptor carboxyl-terminal deletion mutants. *Oncotarget* **6,** 8839–8850 (2015).

276. Vivanco, I. *et al.* Differential sensitivity of glioma- versus lung cancer-specific EGFR mutations to EGFR kinase inhibitors. *Cancer Discov.* **2,** 458–471 (2012).

277. Jackman, D. M. *et al.* Exon 19 Deletion Mutations of Epidermal Growth Factor Receptor Are Associated with Prolonged Survival in Non-Small Cell Lung Cancer Patients Treated with Gefitinib or Erlotinib. *Clin. Cancer Res.* **12,** 3908 (2006).

278. Yoshikawa, S. *et al.* Structural basis for the altered drug sensitivities of non-small cell lung cancer-associated mutants of human epidermal growth factor receptor. *Oncogene* **32,** 27–38 (2013).

279. Yasuda, H. *et al.* Structural, Biochemical, and Clinical Characterization of

Epidermal Growth Factor Receptor (EGFR) Exon 20 Insertion Mutations in Lung Cancer. *Sci. Transl. Med.* **6,** 225er1–225er1 (2014).

280. Shan, Y. *et al.* Oncogenic mutations counteract intrinsic disorder in the EGFR kinase and promote receptor dimerization. *Cell* **149,** 860–870 (2012).

281. Sutto, L. & Gervasio, F. L. Effects of oncogenic mutations on the conformational free-energy landscape of EGFR kinase. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 10616–21 (2013).

282. Red Brewer, M. *et al.* Mechanism for activation of mutated epidermal growth factor receptors in lung cancer. *Proc. Natl. Acad. Sci. U. S. A.* **110,** E3595–604 (2013).

283. Wang, Y. *et al.* Regulation of EGFR nanocluster formation by ionic protein-lipid interaction. *Cell Res.* **24,** 959–76 (2014).

284. Cho, J. *et al.* Cetuximab response of lung cancer-derived EGF receptor mutants is associated with asymmetric dimerization. *Cancer Res.* **73,** 6770–6779 (2013).

285. Jia, Y. *et al.* Overcoming EGFR(T790M) and EGFR(C797S) resistance with mutant-selective allosteric inhibitors. *Nature* **534,** 129–132 (2016).

286. Kobayashi, S. *et al.* EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* **352,** 786–792 (2005).

287. Pao, W. *et al.* Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. *PLoS Med.* **2,** 0225–0235 (2005).

288. Eck, M. J. & Yun, C.-H. Structural and mechanistic underpinnings of the differential drug sensitivity of EGFR mutations in non-small cell lung cancer. *Biochim. Biophys. Acta* **1804,** 559–566 (2010).

289. Ruan, Z. & Kannan, N. Mechanistic Insights into R776H Mediated Activation of Epidermal Growth Factor Receptor Kinase. *Biochemistry* **54,** 4216–4225 (2015).

290. Rusnak, D. W. *et al.* The Effects of the Novel, Reversible Epidermal Growth Factor Receptor/ErbB-2 Tyrosine Kinase Inhibitor, GW2016, on the Growth of Human Normal and Tumor-derived Cell Lines in Vitro and in Vivo. *Mol. Cancer Ther.* **1,** 85–94 (2001).

291. Ethier, S. P., Langton, B. C. & Dilts, C. A. Growth Factor-Independent Proliferation of Rat Mammary Carcinoma Cells by Autocrine Secretion of neu-Differentiation Factor / Heregulin and Transforming Growth Factor-a. *Mol. Carcinog.* **15,** 134–143 (1996).

292. Mullet, A. *et al.* The enlargement of the hormone immune deprivation concept to the blocking of TGF a -autocrine loop : EGFR signaling inhibition. *Cancer Immunol Immunother* **55,** 628–638 (2006).

293. Pao, W. & Girard, N. New driver mutations in non-small-cell lung cancer. *Lancet Oncol.* **12,** 175–180 (2011).

294. Kobayashi, S. *et al.* An alternative inhibitor overcomes resistance caused by a mutation of the epidermal growth factor receptor. *Cancer Res.* **65,** 7096–7101 (2005).

295. Davis, M. I. *et al.* Comprehensive analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **29,** 1046–1051 (2011).

296. Kartalov, E. P., Unger, M. A. & Quake, S. R. Polyelectrolyte surface interface for single-molecule fluorescence studies of DNA polymerase. *Biotechniques* **34,** 505–510 (2003).

297. Zaczek, A., Brandt, B. & Bielawski, K. P. The diverse signaling network of EGFR, HER2, HER3 and HER4 tyrosine kinase receptors and the consequences for therapeutic approaches. *Histol. Histopathol.* **20,** 1005–1015 (2005).

298. Velpula, K. K., Dasari, V. R., Asuthkar, S., Gorantla, B. & Tsung, A. J. EGFR and c-Met Cross Talk in Glioblastoma and Its Regulation by Human Cord Blood Stem Cells. *Transl. Oncol.* **5,** 379–392 (2012).

299. Wang, Z. Transactivation of epidermal growth factor receptor by g protein-coupled receptors: Recent progress, challenges and future research. *Int. J. Mol. Sci.* **17,** 1–12 (2016).

300. Drakaki, A. & Hurvitz, S. A. HER2-Positive Breast Cancer : Update on New and Emerging Agents. *Am. J. Hematol. Oncol.* **11,** 17–23 (2015).

301. Bose, R. *et al.* Activating HER2 mutations in HER2 gene amplification negative breast cancer. *Cancer Discov.* **3,** 224–237 (2013).

302. Wang, Y. *et al.* Different mechanisms for resistance to trastuzumab versus lapatinib in HER2- positive breast cancers – role of estrogen receptor and HER2 reactivation. *Breast Cancer Res.* **13,** 1–19 (2011).

303. D'Amato, V. *et al.* Mechanisms of lapatinib resistance in HER2-driven breast cancer. *Cancer Treat. Rev.* **41,** 877–883 (2015).